AD_____

Award Number: W81XWH-10-1-0173

TITLE: Tissue and Metabolomic Biomarkers of Recurrent Renal Cell Carcinoma

PRINCIPAL INVESTIGATOR: Richard R. Drake (Partnering PI: Alexander Parker)

CONTRACTING ORGANIZATION: Medical University of South Carolina
Charleston, South Carolina 29425

REPORT DATE: July 2014

TYPE OF REPORT: Final Report

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
July 2014	Final Report	1Apr2010 - 30Apr2014

4. TITLE AND SUBTITLE
Tissue and Metabolomic Biomarkers of Recurrent Renal Cell Carcinoma

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-10-1-0173

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Richard R. Drake, Ph.D.

E-Mail: draker@musc.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Medical University of South Carolina
179 Ashley Avenue
Charleston, SC 29425-8908

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The purpose of the study was to harness cutting-edge metabolomic and proteomic biomarker discovery technologies to identify novel biomarkers for ccRCC aggressiveness in primary tumor samples excised during surgery. Fresh-frozen tissue samples from 25 intermediate risk ccRCC patients who experienced progression to metastasis within 3 years of surgery and 25 intermediate risk ccRCC patients who remain progression free after 5 years of follow-up were be evaluated by MALDI mass spectrometry based tissue imaging and metabolomic profiling at Metabolon, Inc. This data has also been extended using new analysis workflows developed at MUSC using the high resolution MALDI-FTICR instrument to analyze lipid and glycan species directly on tissue. An emphasis on examining the molecular changes at the tumor margin interface has also been implemented. These latter methods are currently unique to MUSC and were developed specifically for this project. The cumulative data is currently being analyzed with the clinical data in the partnering PIs facility.

15. SUBJECT TERMS
Kidney cancer, biomarkers, metabolomics, MALDI imaging

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	60	19b. TELEPHONE NUMBER *(include area code)*

Table of Contents

Page

W81XWH-10-1-0173: Tissue and Metabolomic Biomarkers of Recurrent Renal Cell Carcinoma

Partnering Investigators: Richard R. Drake, Ph.D. and Alexander S. Parker, Ph.D.

Final Report from April 1, 2010 to March 30, 2014 for Drake at MUSC

Note: This project was initiated at Easter Virginia Medical School (EVMS). In July 2011, Dr. Drake transferred to the Medical University of South Carolina (MUSC). The time period from June 30, 2011 to March 30, 2012 was not active due to the grant transfer process.

Introduction and Goals:

The incidence and mortality rates for renal cell carcinoma (RCC) have risen steadily for more than 30 years, with a poor 5-year survival rate and a characteristically unpredictable clinical course for the most common clear cell form (ccRCC). The primary treatment for patients with localized ccRCC is surgical excision, which can be highly effective for early stage cancers. However, due to lack of any early detection strategies, approximately 35-40% of patients with no evidence of metastasis at the time of surgery will subsequently experience metastatic progression. Two key clinical issues are the need to 1) identify ways of more accurately predicting which patients will experience metastatic progression following surgery for localized ccRCC and 2) develop new treatments that can be used in combination with surgical excision to reduce progression. The overall goal of our proposed study is to improve our understanding of the underlying mechanisms of clear cell RCC progression and enhance the ability to accurately predict which patients are at greatest risk of progression following surgery. We hypothesize that identification of specific tumor associated proteins directly in histopathological specimens and their corresponding metabolite profile can be linked with our existing panel of biomarkers of clear cell RCC aggressiveness to develop a novel biomarker-based prognostic nomogram/scoring system that can significantly improve the ability to accurately identify individuals most at risk of ccRCC progression following surgery. Four experimental Specific Aims are proposed as follows: **1.** To harness cutting-edge metabolomic and proteomic biomarker discovery technologies to identify novel biomarkers for ccRCC aggressiveness in primary tumor samples excised during surgery; **2.** Combine novel biomarkers from SA1 with existing panel of seven previously published biomarkers of ccRCC aggressiveness to develop composite biomarker-based algorithm for predicting progression following surgery for ccRCC; **3.** To harness cutting-edge metabolomic and proteomic biomarker discovery technologies to identify novel biomarkers that are differentially expressed in paired samples of primary and metastatic ccRCC; and **4.** To independently validate the differential expression of the candidate biomarkers identified in SA3 and estimate the association of the expression of these biomarkers in metastatic ccRCC with time to death. To accomplish this, fresh-frozen tissue samples from 25 intermediate risk ccRCC patients who experienced progression to metastasis within 3 years of surgery and 25 intermediate risk ccRCC patients who remain progression free after 5 years of follow-up will be evaluated by MALDI mass spectrometry based tissue imaging and metabolomic profiling. Also, the same tissue imaging and metabolomic approaches will be applied to fresh-frozen tissue samples from 15 patients with matched primary ccRCC tumor and metastatic lung ccRCC tumor pairs. A

novel biomarker-based scoring algorithm for predicting ccRCC progression using a cohort of 1,500 patients undergoing nephrectomy for localized ccRCC will also be developed. An additional 250 patients who have archived tumor blocks available from both primary and metastatic ccRCC will also be evaluated for development of a biomarker panel. For impact, the identification of molecular biomarkers within tumor tissue that correlate with risk of ccRCC progression has the potential to not only improve prognostic assessment and enhance post-operative surveillance, but also to inform on the underlying biology of ccRCC aggressiveness as well as to provide rational targets and strategies for therapeutic intervention.

Keywords:
Clear cell renal cell carcinoma, biomarkers, glycoproteins, proteomics, metabolomics, lipidomics, glycomics, mass spectrometry, MALDI imaging mass spectrometry

Overall Project Summary:
In the Statement of Work for this project, 7 Tasks were proposed over a three year period. The first three Tasks were proposed to be completed within the first 18 months of the project. Overall, excellent progress has been made on these Tasks, and no changes in the last Statement of Work (revised with the move to MUSC) and/or research directions are expected. The scientific progress for completion of three tasks involving the Drake laboratory are summarized, followed by a data example summary in the Appendix Materials.

Task 1. **Identify differentially expressed proteomic biomarkers by MALDI mass spectrometry imaging in a cohort of patients with localized ccRCC classified as being at intermediate risk for recurrence (i.e. progression to metastasis) following surgery.**

a. Perform MALDI-TOF MS imaging analysis of 50 ccRCC tissues classified as intermediate risk for recurrence; 25 that experienced recurrence within 3 years of surgery, and 25 that remain recurrence free > 5 years following surgery.
b. Analyze peak intensity data inter- and intra-sample to determine differentially expressed proteins and peptides.

MILESTONE: Establish a panel of 5-10 differentially expressed peak markers that distinguish intermediate risk ccRCC patients who experience progression to metastasis following surgery from intermediate risk ccRCC patients who do not experience progression to metastasis. Months 1-18 (Drake)

Experimental Progress: This task was initiated by Dr. Parker with the selection of 50 frozen ccRCC samples from the biorepository at Mayo Clinic. Samples from patients with non-recurrent disease were selected from individuals with no evidence of disease 5 years after surgery for ccRCC. Samples from patients with recurrent disease were selected from individuals who had detectable metatstatic ccRCC within three years of primary nephrectomy. In addition, samples were matched in pairs to age, race, gender, and pathology information. These samples were dehydrated in ethanol and sprayed with CHCA matrix, followed by tissue imaging in the UltraFlex III MALDI-TOF/TOF instrument (Bruker Daltonics). Following reading of each slide by a pathologist (Dr. Dean Troyer, at EVMS), regions of non-necrotic tumor were selected as

regions of interest (ROIs) for selection of spectra. On average, 50-100 individual spectra per ROI were selected. As summarized from the original Year 1 report, usable spectra for 22 recurrent ccRCC and 26 non-recurrent ccRCC samples were obtained. These procedures and experiments cumulatively took approximately 10 months to complete, and represent a total of 48 ccRCC and 10 normal matched renal tissues. Since June 2011, the data analysis of the obtained ROI spectra for each sample was done using FlexAnalysis and FlexImaging software (from Bruker Daltonics). Tandem mass spectrometry of 10 tissues was also done to correlate protein peaks identified in the MALDI imaging data to small proteins/peptides. This has facilitated identification of thymosin beta 4, thymosin beta 10, S100 A8, S100A9, S100A11 as clear markers distinguishing tumor from normal tissues. The proteomic analyses were essentially completed for the protein analysis aspects in the laboratory by the end of Year 2, and the data was published (see Appendix)

The data analysis continues with Dr. Parker for other molecular correlates (as in Task 3).

Task 2. **Identify differentially expressed metabolite biomarkers in the same ccRCC tissue intermediate risk cohort described in Task 1.**

a. Perform metabolite analysis of 50 ccRCC tissues classified as intermediate risk for recurrence; 25 that experienced a recurrence within 3 years of surgery, and 25 that remain recurrence free > 5years following surgery.
b. Determine any differentially expressed metabolites reflective of disease recurrence and/or non-recurrence.

MILESTONE: Determine a panel of ccRCC metabolites (15-20) to predict disease recurrence at the time of surgery. Months 9-15 (Metabolon, Parker)

Experimental Progress: For this task, metabolite analysis is contracted to Metabolon, Inc, which provided initial requirements for tissue preparation. As indicated in Task 1, the tissue slices for analysis of metabolites were prepared at the time of MALDI slide preparation. Tumor tissues (approx 25 mg/sample) were sent frozen to Metabolon in month 9 of the project, and the negotiated analysis costs already paid to the company. Dr. Parker and Dr. Drake were subsequently informed by Metabolon that at least 50 mg, and preferably 100 mg, of tissue from each sample will give the best results. Subsequently, larger amounts of RCC tissue samples were sent to Metabolon in the Fall of 2012 and analysis was completed in December 2012. The analysis report was attached with Dr. Parker's last annual report, and is not included in this report. In the past few months this summer (June/July 2013), a subset of the tissues analyzed by Metabolon have been examined by MALDI imaging, with an emphasis on determining the distribution of differentially expressed phospholipid and small metabolite biomolecules from the Metabolon report. Examples are provided in the Appendix.

Task 3. Develop a composite biomarker-based algorithm for predicting progression following surgery for ccRCC.

a. Examine the co-expression of an existing seven-member biomarker panel (*survivin, B7-H1, B7-H4, Ki-67, IGF-IR, IMP-3 and CA-IX*) from a cohort of over 1500 paraffin-embedded formalin-fixed

samples to create a BioMarker prediction algorithm.

b. Assess the collinearity between all biomarkers using multivariable Cox proportional hazards regression to evaluate if the full panel of biomarkers are independent predictors of time to ccRCC progression.

c. Evaluate the improvement in outcome prediction for BioMarker over the individual biomarkers that were combined to create it.

d. Evaluate whether MALDI-MS tissue biomarker and metabolite candidates from Tasks 1 and 2 improve the outcome prediction for BioMarker for future validation studies.

MILESTONE: Develop a BioMarker classification algorithm for predicting progression following surgery for ccRCC. Months 1-18 (Parker)

From this task we have develop a scoring algorithm called BioPROG that is being currently developed as a clinically available test at Mayo Clinic in Florida through the Biomarker Discovery Program within the Mayo Center for Individualized Medicine. This test will help urologic surgeons to better distinguish between those ccRCC patients who will progress and die after surgery form those with a more indolent course. This opens up opportunities clinically to 1) target more aggressive care and surveillance to those that need it, 2) avoid unnecessary imaging modalities for those that truly are at low risk of recurrence and design better trials for new targeted therapeutics in the future by more accurately idnetifyingt hsoe with moderate and high risk disease for these trials.

To create BioPROG we evaluated the protein expression of the 7 genes in our original panel proposed in the grant (as well as new candidates that arose after the grant initiation through a collaboration with Dr. Jim Brugarlos at UT Southwestern) to develop a score for predicting ccRCC progression (metastasis or RCC death) in patients with clinically localized ccRCC (M0 disease).

First off, **Table 1** provides characteristics of the patients from our overall cohort that were included in the analyses and those that were not (i.e. because tissue was missing, etc.). These analyses help us determine if the excluding of patients has introduced any potential biases.

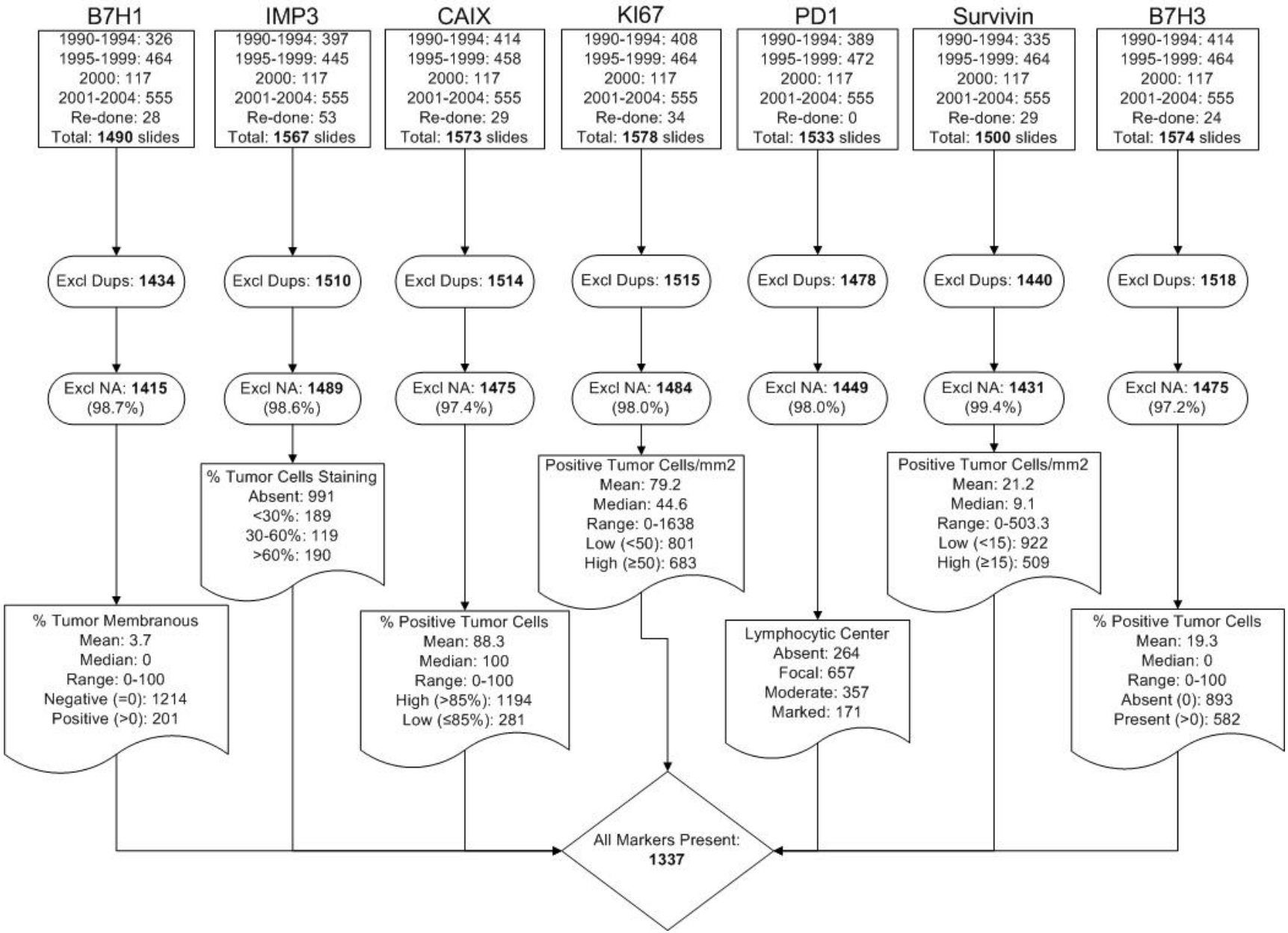
Table 1: Clinical and pathological information for BioProg Analyses

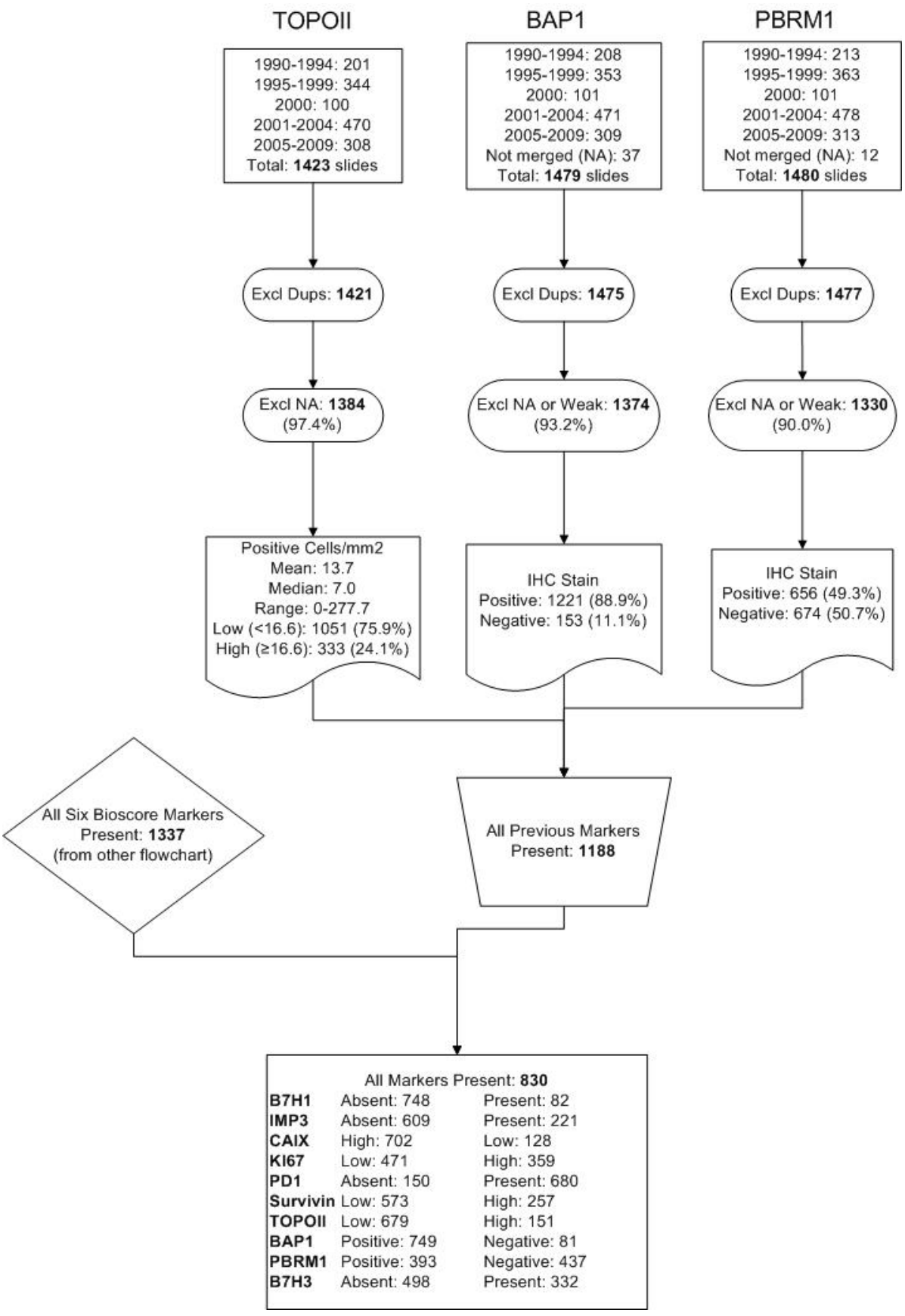
	Not-included (N=827)	BioProg (N=830)	Total (N=1657)	p value
Male				0.0510
No	311 (37.6%)	274 (33.0%)	585 (35.3%)	
Yes	516 (62.4%)	556 (67.0%)	1072 (64.7%)	
Age at Surgery				0.7033
Mean	63.5	63.3	63.4	
Median	64.6	64.2	64.5	
Range	(19.8-90.2)	(22.2-90.0)	(19.8-90.2)	
Tumor Size				0.8175
Mean	6.1	6.2	6.1	
Median	5.5	5.0	5.2	

Table 1: Clinical and pathological information for BioProg Analyses

	Not-included (N=827)	BioProg (N=830)	Total (N=1657)	p value
Range	(0.5-29.0)	(1.0-22.0)	(0.5-29.0)	
TNM Stage				0.2174
Missing	2	4	6	
1	479 (58.1%)	509 (61.6%)	988 (59.8%)	
2	110 (13.3%)	118 (14.3%)	228 (13.8%)	
3	229 (27.8%)	192 (23.2%)	421 (25.5%)	
4	7 (0.8%)	7 (0.8%)	14 (0.8%)	
Nuclear Grade				0.0015
1	65 (7.9%)	65 (7.8%)	130 (7.8%)	
2	376 (45.5%)	376 (45.3%)	752 (45.4%)	
3	311 (37.6%)	352 (42.4%)	663 (40.0%)	
4	75 (9.1%)	37 (4.5%)	112 (6.8%)	
Coagulative Tumor Necrosis				0.0130
No	609 (73.6%)	655 (78.9%)	1264 (76.3%)	
Yes	218 (26.4%)	175 (21.1%)	393 (23.7%)	
SSIGN Score				0.0522
Missing	4	5	9	
0	263 (32.0%)	263 (31.9%)	526 (31.9%)	
1	52 (6.3%)	79 (9.6%)	131 (7.9%)	
2	94 (11.4%)	98 (11.9%)	192 (11.7%)	
3	94 (11.4%)	100 (12.1%)	194 (11.8%)	
4	72 (8.7%)	69 (8.4%)	141 (8.6%)	
5	70 (8.5%)	80 (9.7%)	150 (9.1%)	
6	31 (3.8%)	31 (3.8%)	62 (3.8%)	
7	80 (9.7%)	66 (8.0%)	146 (8.9%)	
8	5 (0.6%)	8 (1.0%)	13 (0.8%)	
9	49 (6.0%)	26 (3.2%)	75 (4.6%)	
10	1 (0.1%)	0 (0.0%)	1 (0.1%)	
11	12 (1.5%)	5 (0.6%)	17 (1.0%)	
(report generated on 26MAR2014)				

The figures that follow below show an overall schematic of the patients that made it in to our analyses for each biomarker:


B7H1
1990-1994: 326
1995-1999: 464
2000: 117
2001-2004: 555
Re-done: 28
Total: 1490 slides
Excl Dups: 1434
Excl NA: 1415
(98.7%)
% Tumor Membranous
Mean: 3.7
Median: 0
Range: 0-100
Negative (=0): 1214
Positive (>0): 201
IMP3
1990-1994: 397
1995-1999: 445
2000: 117
2001-2004: 555
Re-done: 53
Total: 1567 slides
Excl Dups: 1510
Excl NA: 1489
(98.6%)
% Tumor Cells Staining
Absent: 991
<30%: 189
30-60%: 119
>60%: 190
CAIX
1990-1994: 414
1995-1999: 458
2000: 117
2001-2004: 555
Re-done: 29
Total: 1573 slides
Excl Dups: 1514
Excl NA: 1475
(97.4%)
% Positive Tumor Cells
Mean: 88.3
Median: 100
Range: 0-100
High (>85%): 1194
Low (≤85%): 281
KI67
1990-1994: 408
1995-1999: 464
2000: 117
2001-2004: 555
Re-done: 34
Total: 1578 slides
Excl Dups: 1515
Excl NA: 1484
(98.0%)
Positive Tumor Cells/mm2
Mean: 79.2
Median: 44.6
Range: 0-1638
Low (<50): 801
High (≥50): 683
PD1
1990-1994: 389
1995-1999: 472
2000: 117
2001-2004: 555
Re-done: 0
Total: 1533 slides
Excl Dups: 1478
Excl NA: 1449
(98.0%)
Lymphocytic Center
Absent: 264
Focal: 657
Moderate: 357
Marked: 171
Survivin
1990-1994: 335
1995-1999: 464
2000: 117
2001-2004: 555
Re-done: 29
Total: 1500 slides
Excl Dups: 1440
Excl NA: 1431
(99.4%)
Positive Tumor Cells/mm2
Mean: 21.2
Median: 9.1
Range: 0-503.3
Low (<15): 922
High (≥15): 509
B7H3
1990-1994: 414
1995-1999: 464
2000: 117
2001-2004: 555
Re-done: 24
Total: 1574 slides
Excl Dups: 1518
Excl NA: 1475
(97.2%)
% Positive Tumor Cells
Mean: 19.3
Median: 0
Range: 0-100
Absent (0): 893
Present (>0): 582
All Markers Present:
1337


TOPOII
1990-1994: 201
1995-1999: 344
2000: 100
2001-2004: 470
2005-2009: 308
Total: 1423 slides
Excl Dups: 1421
Excl NA: 1384
(97.4%)
Positive Cells/mm2
Mean: 13.7
Median: 7.0
Range: 0-277.7
Low (<16.6): 1051 (75.9%)
High (≥16.6): 333 (24.1%)
BAP1
1990-1994: 208
1995-1999: 353
2000: 101
2001-2004: 471
2005-2009: 309
Not merged (NA): 37
Total: 1479 slides
Excl Dups: 1475
Excl NA or Weak: 1374
(93.2%)
IHC Stain
Positive: 1221 (88.9%)
Negative: 153 (11.1%)
PBRM1
1990-1994: 213
1995-1999: 363
2000: 101
2001-2004: 478
2005-2009: 313
Not merged (NA): 12
Total: 1480 slides
Excl Dups: 1477
Excl NA or Weak: 1330
(90.0%)
IHC Stain
Positive: 656 (49.3%)
Negative: 674 (50.7%)
All Six Bioscore Markers
Present: 1337
(from other flowchart)
All Previous Markers
Present: 1188
All Markers Present: 830
B7H1 Absent: 748 Present: 82
IMP3 Absent: 609 Present: 221
CAIX High: 702 Low: 128
KI67 Low: 471 High: 359
PD1 Absent: 150 Present: 680
Survivin Low: 573 High: 257
TOPOII Low: 679 High: 151
BAP1 Positive: 749 Negative: 81
PBRM1 Positive: 393 Negative: 437
B7H3 Absent: 498 Present: 332

Below we show the univariate Cox models for each individual marker in the first table and then compare them to the original publications for each marker in the next Table:

Univariate Cox Models:

	Bioscore 2 (N=830)		RCC Specific Survival (167 events)		Risk of Metastases (216 events)		Progression-free Survival (225 events)	
Gene	Low Risk	High Risk	Age Adjusted HR (95% CI)	P-value	Age Adjusted HR (95% CI)	P-value	Age Adjusted HR (95% CI)	P-value
B7H1	Absent: 748	Present: 82	3.639 (2.540 - 5.214)	1.93E-12	3.451 (2.482 - 4.798)	1.78E-13	3.369 (2.434 - 4.664)	2.48E-13
IMP3	Absent: 609	Present: 221	5.892 (4.305 - 8.064)	<2.0E-16	4.883 (3.725 - 6.400)	<2.0E-16	4.711 (3.615 - 6.139)	<2.0E-16
CAIX	High: 702	Low: 128	1.647 (1.150 - 2.361)	0.00655	1.468 (1.057 - 2.040)	0.022	1.431 (1.035 - 1.980)	0.0303
KI67	Low: 471	High: 359	2.867 (2.083 - 3.946)	1.01E-10	2.531 (1.922 - 3.334)	3.86E-11	2.560 (1.953 - 3.354)	9.35E-12
PD1	Absent: 150	Present: 680	1.472 (0.931 - 2.328)	0.098	2.246 (1.417 - 3.561)	0.000576	2.109 (1.357 - 3.276)	0.000903
Survivin	Low: 573	High: 257	4.884 (3.557 - 6.705)	<2.0E-16	4.259 (3.242 - 5.596)	<2.0E-16	4.014 (3.075 - 5.240)	<2.0E-16
TOPOII	Low: 679	High: 151	3.177 (2.314 - 4.360)	8.50E-13	2.445 (1.827 - 3.271)	1.78E-09	2.356 (1.768 - 3.141)	5.07E-09
BAP1	Positive: 749	Negative: 81	2.400 (1.615 - 3.566)	1.48E-05	1.650 (1.115 - 2.441)	0.0122	1.764 (1.213 - 2.565)	0.003
PBRM1	Positive: 393	Negative: 437	1.163 (0.850 - 1.592)	0.346	1.532 (1.155 - 2.032)	0.00305	1.419 (1.079 - 1.868)	0.0124
B7H3	Absent: 498	Present: 332	2.646 (1.932 - 3.625)	1.33E-09	2.069 (1.577 - 2.714)	1.52E-07	2.056 (1.576 - 2.683)	1.10E-07
Age only	--	--	1.021 (1.007 - 1.035)	0.0025	1.0076 (0.996 - 1.019)	0.20	1.0089 (0.998 - 1.020)	0.13

Comparison with Original Papers:

Gene	Paper	N	Analysis (closest to Bioscore 2)	HR	P-value
B7H1	Cancer Research 2006	268 (breakdown unclear)	RCC death in M0	4.13 (2.49 - 6.86)	<0.001
IMP3	Cancer 2008	Absent: 466 Present: 163	Risk of metastases in M0 (161 events)	4.71 (2.44 - 6.43)	<0.001
CAIX	JCO 2007	High: 567 Low: 163	Age adjusted RCC death in M0 and M1 (241 events)	1.65 (1.25 - 2.18)	<0.001
KI67	Cancer 2007	Low: 460 High: 281	RCC death in M0 and M1 (238 events)	2.36 (1.80 - 3.09)	<0.001
PD1	CCR 2007	Absent: 190 Present: 77	RCC death in M0 and M1 (52 events)	2.24 (1.30 - 3.86)	0.004
Survivin	Cancer 2006	Low: 203 High: 79	RCC death in M0	5.1 (3.1 - 8.4)	<0.001
TOPOII	EU 2014	Low: 1046 High: 332	Age adjusted RCC death in M0 (244 events)	2.73 (2.12 - 3.54)	1.79E-14
BAP1	Cancer 2014	Positive: 1196 Negative: 148	Age adjusted RCC death in M0 (252 events)	3.06 (2.28 - 4.10)	6.77E-14
PBRM1	CCR 2014?	Positive: 656 Negative: 674	Age adjusted risk of metastases in M0 (330 events)	1.46 (1.16 - 1.83)	0.0011
B7H3	OncoTargets and Therapy	---	---	---	---

Below we provide tables showing all 10 markers in the same Cox model. First for the outcome of RCC –specific survival, then time to metastasis and then Progression Free survival.

Cox Models with 10 Genes Present

	Age Adjusted RCC-SS				
variable	**Coefficient**	**HR**	**P-value**	**L95**	**U95**
b7h1	0.382938031	1.466587144	0.054070502	0.993324366	2.165332821
imp3	1.144993938	3.142422308	3.84E-09	2.146970913	4.599418604
caix	0.464819331	1.591726588	0.01449504	1.096554508	2.310503958
ki67	0.047533476	1.048681306	0.808747059	0.71364234	1.541013501
survivin	0.847522872	2.333858419	3.90E-05	1.558409149	3.495163722
topo	0.273556698	1.314631894	0.125182171	0.926720768	1.864916678
bap1	0.228313555	1.256479239	0.323142217	0.79883722	1.976297597
pd1	-0.207235766	0.81282799	0.403002659	0.500106926	1.321096165
pbrm1	0.307973306	1.360664667	0.088520399	0.954644014	1.939370391
b7h3	0.24313766	1.275244162	0.182542624	0.891931138	1.823288373
age at surgery	0.014395138	1.014499247	0.046786974	1.000204626	1.028998163

	Age Adjusted Time to Metastasis				
variable	**Coefficient**	**HR**	**P-value**	**L95**	**U95**
b7h1	0.505353445	1.657571278	0.00551345	1.160065687	2.36843704
imp3	1.084299834	2.957368446	1.22E-10	2.125750541	4.114324779
caix	0.298875162	1.348341288	0.083911193	0.960754632	1.89228776
ki67	-0.013719733	0.986373954	0.935833716	0.706290552	1.377525969
survivin	0.834175755	2.3029151	2.47E-06	1.627606647	3.258415026
topo	0.113369216	1.120045396	0.486597447	0.81382771	1.541483134
bap1	-0.098654761	0.90605546	0.661678165	0.582444365	1.40946766
pd1	0.320254468	1.377478244	0.191440587	0.851937208	2.227213807
pbrm1	0.443514858	1.558174367	0.005078301	1.142574989	2.124943554
b7h3	0.087254352	1.091174187	0.578665786	0.801960601	1.484687782
age at surgery	-0.004665412	0.995345454	0.444883261	0.983503117	1.007330384

	Age Adjusted PFS				
variable	**Coefficient**	**HR**	**P-value**	**L95**	**U95**
b7h1	0.497155528	1.644038192	0.005558008	1.15689127	2.336314264
imp3	1.052090777	2.863632078	1.69E-10	2.073548201	3.954761542
caix	0.282720926	1.326734853	0.096678973	0.950418786	1.852052375
ki67	0.056973878	1.058628156	0.733492897	0.762622653	1.469525681
survivin	0.766029606	2.151208132	9.09E-06	1.53372745	3.017287346
topo	0.08471678	1.088408764	0.598961591	0.793727155	1.492494782
bap1	-0.032596358	0.967929177	0.880179475	0.633548572	1.478792525
pd1	0.273602865	1.314692588	0.243673351	0.829973277	2.082496686
pbrm1	0.381354627	1.464266782	0.013115727	1.083331784	1.979151023
b7h3	0.084633823	1.088318477	0.582706711	0.804708847	1.471882795
age at surgery	-0.002306357	0.997696301	0.699663756	0.986073855	1.009455736

In the following tables we show the pairwise interactions with each biomarker:

Pairwise interactions (PFS) with interaction p-value < 0.1

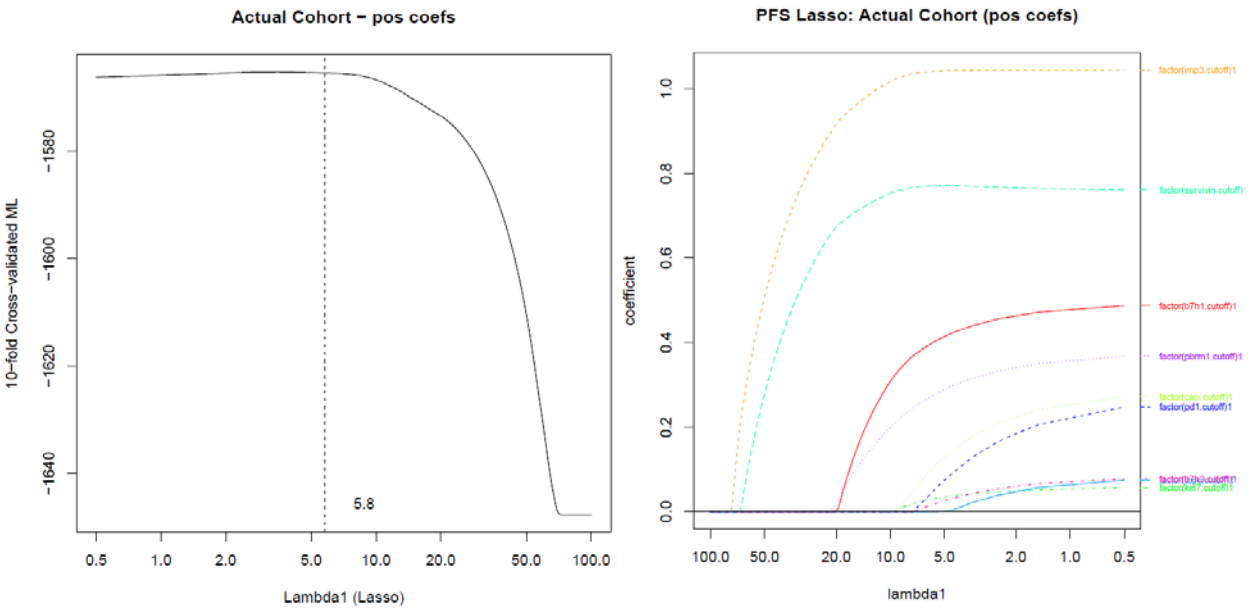
		Marker 1				Marker 2				Interaction			
marker1	marker2	OR	L95	U95	P	OR	L95	U95	P	OR	L95	U95	P
pd1	ki67	4.48	1.95	10.26	0.000395	12.38	4.84	31.7	1.56E-07	0.16	0.06	0.43	0.000272
survivin	pd1	14.78	6.19	35.28	1.30E-09	3.23	1.57	6.66	0.00148	0.22	0.09	0.56	0.00127
bap1	imp3	2.34	1.22	4.51	0.0108	5.46	4.1	7.27	0	0.28	0.13	0.64	0.00221
bap1	pd1	6.43	2.17	19.02	0.000767	2.38	1.47	3.87	0.000458	0.23	0.07	0.73	0.0125
topo	pd1	6.12	2.64	14.18	2.37E-05	2.92	1.62	5.26	0.000358	0.34	0.14	0.82	0.0166
pbrm1	topo	1.63	1.18	2.27	0.00323	3.4	2.19	5.27	5.19E-08	0.52	0.29	0.94	0.0297
topo	caix	2.03	1.45	2.82	2.93E-05	1.16	0.77	1.75	0.475	2.13	1.08	4.2	0.0301
pbrm1	b7h1	1.66	1.22	2.26	0.00142	4.89	3.08	7.76	1.74E-11	0.51	0.27	0.99	0.0466
pbrm1	caix	1.68	1.23	2.3	0.00125	2.05	1.31	3.21	0.0016	0.53	0.27	1.03	0.0596
pd1	imp3	2.45	1.28	4.71	0.00706	9.67	4.18	22.4	1.18E-07	0.43	0.18	1.04	0.061
pbrm1	ki67	1.82	1.18	2.82	0.00713	3.43	2.23	5.28	2.14E-08	0.59	0.34	1.03	0.0616
survivin	b7h1	3.2	2.38	4.31	1.13E-14	0.96	0.35	2.62	0.939	2.76	0.95	8.05	0.0625

Pairwise interactions (PFS) with age at surgery

	Marker				Age at Surgery				Interaction			
marker	OR	L95	U95	P	OR	L95	U95	P	OR	L95	U95	P
b7h1	7.01	1.34	36.67	0.0211	1.01	1	1.02	0.105	0.99	0.96	1.01	0.38
imp3	12.51	2.87	54.51	0.000768	1.01	1	1.03	0.119	0.98	0.96	1.01	0.186
caix	8.79	1.62	47.68	0.0118	1.01	1	1.03	0.0251	0.97	0.95	1	0.0354
ki67	2.39	0.54	10.66	0.253	1.01	0.99	1.02	0.506	1	0.98	1.02	0.928
pd1	3.68	0.32	42.38	0.295	1.02	0.98	1.05	0.398	0.99	0.95	1.03	0.647
survivin	14.37	3.3	62.49	0.000381	1.02	1	1.03	0.0848	0.98	0.96	1	0.0828
b7h3	2.2	0.51	9.59	0.293	1.01	0.99	1.03	0.224	1	0.98	1.02	0.926
topo	4.43	0.89	21.98	0.0683	1.01	1	1.03	0.0884	0.99	0.97	1.02	0.434
bap1	3.12	0.39	24.68	0.281	1.01	1	1.02	0.12	0.99	0.96	1.02	0.586
pbrm1	3.31	0.73	15.02	0.121	1.01	1	1.03	0.158	0.99	0.96	1.01	0.265

To build our final Cox multivariate model that is the basis for BioPROG, we considered three approaches to combine the biomarkers in a single model. The first is based on Lasso bootstrapping, the second is based on Ridge Regression boot strapping and the final on PartDSA analysis branching methods.

Summary of Lasso over 500 Bootstrap Iterations



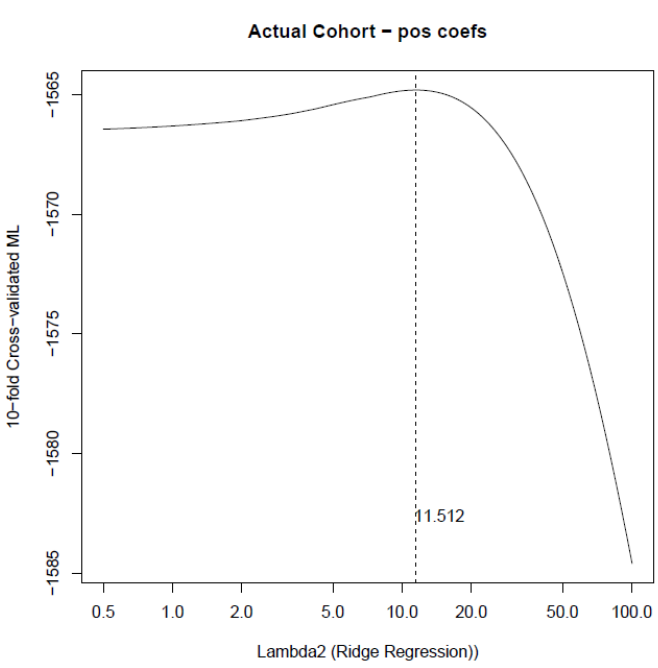
The optimum lambda was found to be 5.8, and that yields the coefficients below:

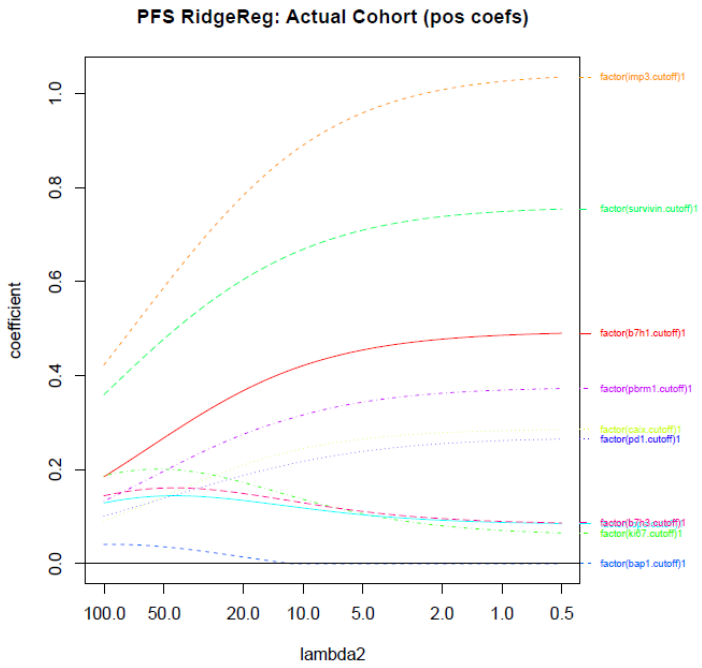
```
imp3:     1.041
survivin: 0.771
b7h1:     0.400
pbrm1:    0.275
caix:     0.104
pd1:      0.048
ki67:     0.030
b7h3:     0.017
topo:     0.000
bap1:     0.000
```

To evaluate the performance of our penalized model, we performed 500 iterations of the bootstrap and evaluated lambda/coefficients/c-index at each.

	# Non-zero	Mean	SD	L95	U95
Lambda1	500	3.0645	1.6983	2.9156	3.2134
b7h1 coef	486	0.4578	0.2079	0.4394	0.4763
imp3 coef	500	1.0226	0.1582	1.0087	1.0364
caix coef	394	0.293	0.1604	0.2772	0.3088
ki67 coef	255	0.1449	0.1035	0.1322	0.1576
survivin coef	500	0.7252	0.1589	0.7113	0.7392
topo2 coef	254	0.1602	0.1214	0.1453	0.1751
bap1 coef	91	0.2127	0.1621	0.1794	0.246
pd1 coef	354	0.2894	0.1962	0.269	0.3099
pbrm1 coef	489	0.3474	0.1496	0.3341	0.3606
b7h3 coef	308	0.1509	0.1107	0.1386	0.1633
bootstrap c-index	500	0.7643	0.0178	0.7628	0.7659
observed c-index	500	0.7574	0.0039	0.7571	0.7578
optimism	500	-0.0069	0.0174	-0.0084	-0.0053

Summary of Ridge Regression over 500 Bootstrap Iterations





The optimum lambda was found to be 11.512, and that yields the coefficients below:

```
imp3:     0.872
survivin: 0.658
b7h1:     0.412
pbrm1:    0.309
caix:     0.239
pd1:      0.212
ki67:     0.144
b7h3:     0.134
topo:     0.122
bap1:     0.000
```

To evaluate the performance of our penalized model, we performed 500 iterations of the bootstrap and evaluated lambda/coefficients/c-index at each.

	# Non-zero	Mean	SD	L95	U95
Lambda2	500	6.4716	3.0551	6.2038	6.7394
b7h1 coef	498	0.4555	0.1845	0.4393	0.4717
imp3 coef	500	0.936	0.1551	0.9224	0.9496
caix coef	480	0.2893	0.1474	0.2761	0.3025
ki67 coef	353	0.1658	0.1048	0.1549	0.1768
survivin coef	500	0.6836	0.141	0.6712	0.696
topo2 coef	371	0.1597	0.1116	0.1484	0.1711
bap1 coef	234	0.1479	0.1205	0.1324	0.1633
pd1 coef	459	0.2749	0.1656	0.2597	0.29
pbrm1 coef	498	0.3592	0.1333	0.3475	0.3709
b7h3 coef	403	0.1644	0.1051	0.1541	0.1746
bootstrap c-index	500	0.7646	0.0176	0.7631	0.7662
observed c-index	500	0.7581	0.0031	0.7578	0.7583
optimism	500	-0.0066	0.0174	-0.0081	-0.005

Lasso including all pairwise interactions

To evaluate which interactions were important in these models, we included all pairwise interactions in a lasso model (with and without the positive coefficient constraint, since interactions might require a negative interaction).

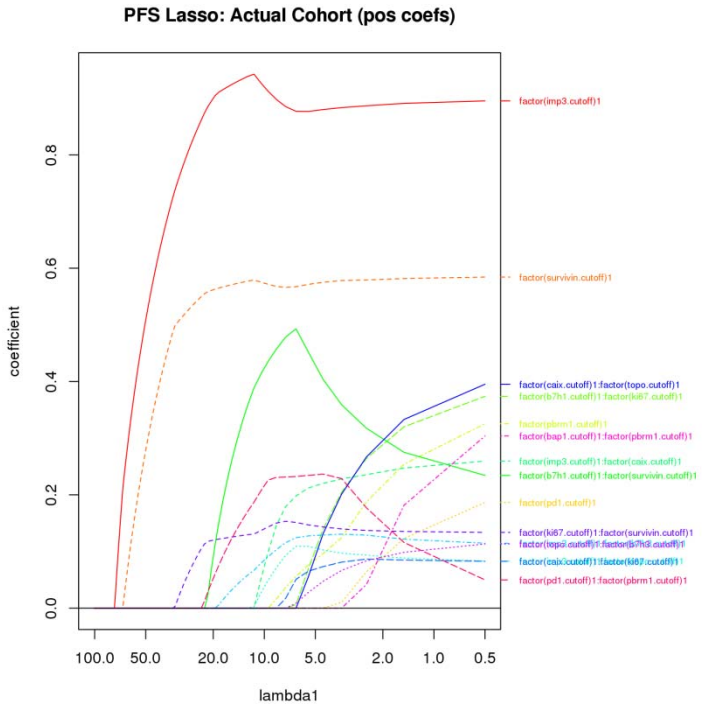
Without the positive constraint, the optimum lambda1 was 6.964. Coefficients:

imp3	0.880
survivin	0.566
pbrm1	0.047
b7h1:survivin	0.489
pd1:pbrm1	0.232
imp3:caix	0.189
ki67:survivin	0.153
imp3:b7h3	0.122
imp3:topo	0.104
caix:ki67	0.036
topo:b7h3	0.001

With positive constraint, the optimum lambda1 was 6.697. Coefficients:

imp3	0.878
survivin	0.567
pbrm1	0.052
b7h1:survivin	0.495
pd1:pbrm1	0.232
imp3:caix	0.193
ki67:survivin	0.152
imp3:b7h3	0.123
imp3:topo	0.107
caix:ki67	0.045
topo:b7h3	0.005

Since the models were essentially equal, we only include the positive coefficients plot.



Evaluating Univariate Cutoffs Generated via partDSA (for numeric biomarkers)

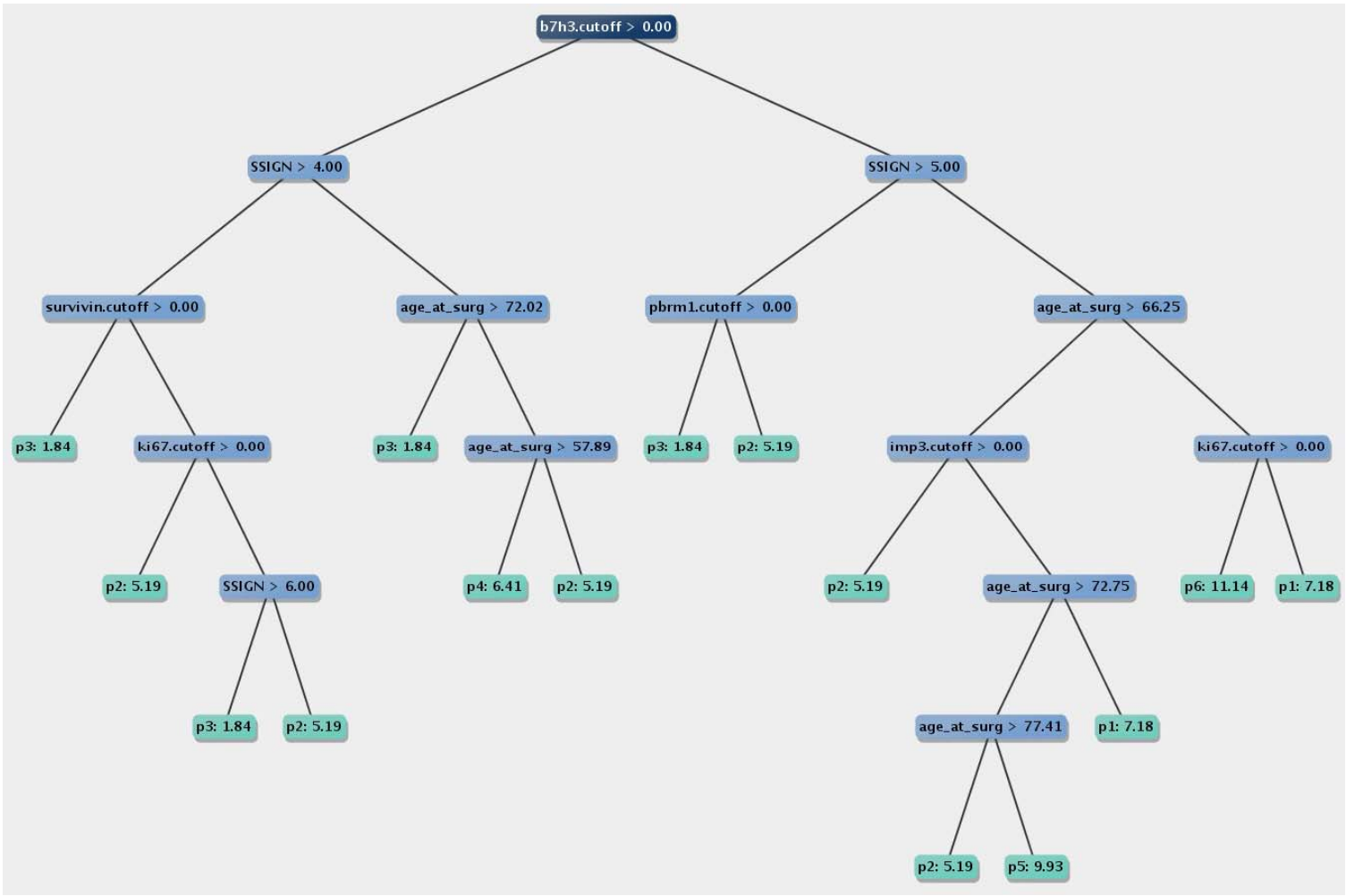
	Cutoffs (low-risk group)		PFS with Established Cutoffs (225 events)		PFS with partDSA Cutoffs (225 events)	
Gene	Established	partDSA	Age Adjusted HR (95% CI)	P-value	Age Adjusted HR (95% CI)	P-value
B7H1	=0	≤ 5	3.369 (2.434 - 4.664)	2.48E-13	3.083 (2.001 - 4.751)	3.33E-07
CAIX	> 85	> 50	1.431 (1.035 - 1.980)	0.0303	1.464 (0.950 - 2.257)	0.084
KI67	< 50	< 90.55	2.560 (1.953 - 3.354)	9.35E-12	2.619 (2.000 - 3.430)	2.61E-12
Survivin	< 15	≤ 20.3	4.014 (3.075 - 5.240)	<2.0E-16	4.167 (3.189 - 5.445)	<2.0E-16
TOPOII	< 16.6	≤ 33.2	2.356 (1.768 - 3.141)	5.07E-09	3.807 (2.626 - 5.519)	1.72E-12
B7H3	=0	=0	2.056 (1.576 - 2.683)	1.10E-07	2.056 (1.576 - 2.683)	1.10E-07

SSIGN Cutoffs via partDSA (same cutoffs whether n=830 or n=1657 cohort)

	Cutoffs (n=825)		PFS with Established Cutoffs (224 events)		PFS with partDSA Cutoffs (224 events)	
SSIGN	Established	partDSA	Age Adjusted HR (95% CI)	P-value	Age Adjusted HR (95% CI)	P-value
Low	0-3 (540)	0-4 (609)	Ref	NA	Ref	NA
Med	4-7 (246)	5-6 (111)	7.74 (5.65 - 10.61)	<2.0E-16	5.86 (4.18 - 8.22)	<2.0E-16
High	8+ (39)	7+ (105)	22.44 (14.45 - 34.86)	<2.0E-16	13.94 (10.18 - 19.09)	<2.0E-16

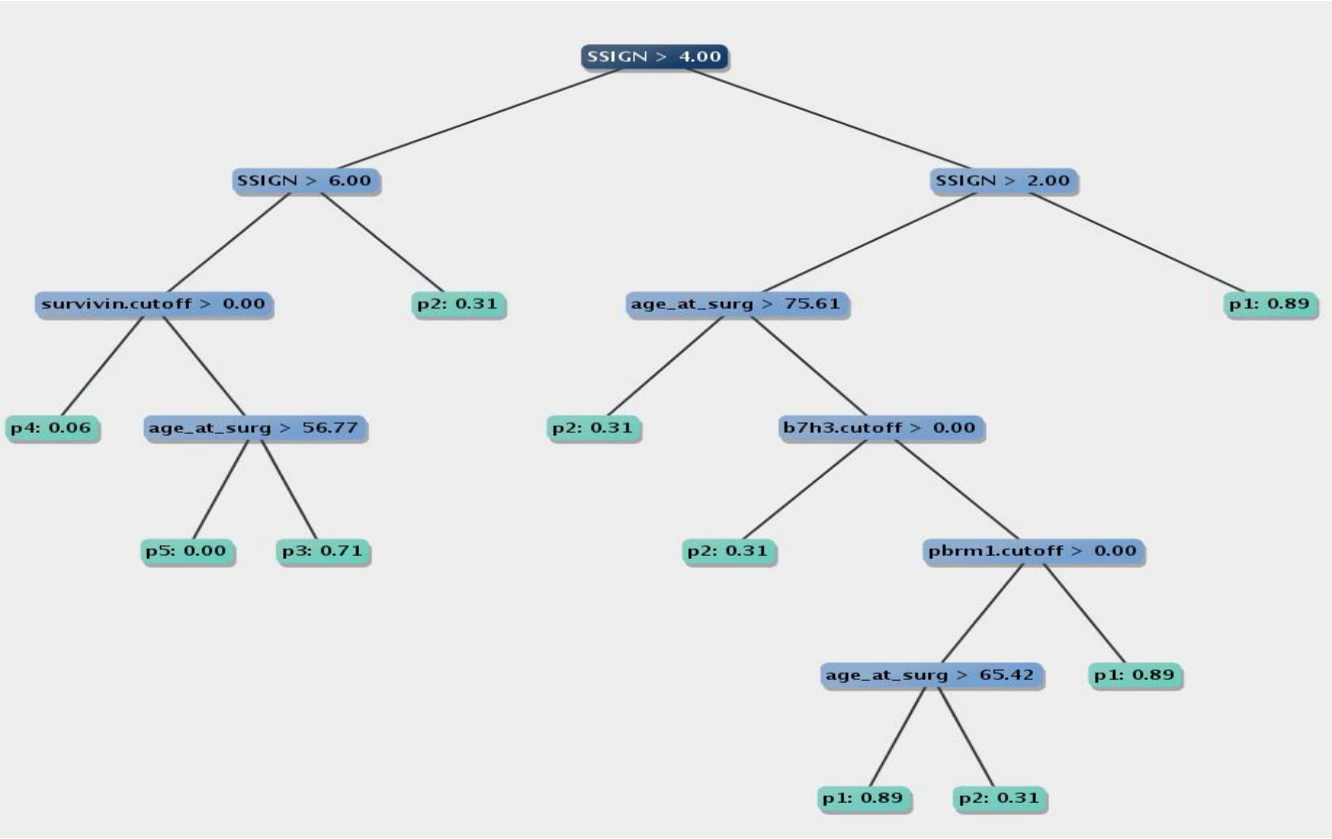
partDSA output with age SSIGN/10-markers - IPCW weighting scheme

```
# partitions   mean CV error   sd CV error   test risk
1              19.908413       7.479193      22.804905
2              21.066529       11.771594     15.641503
3              20.620430       12.126912     13.641046
4              20.662944       11.926582     13.704541
5              21.183672       11.839879     13.509006
6              20.076213       10.606785     12.241115


Variable importance matrix:
                 COG=1 COG=2 COG=3 COG=4 COG=5 COG=6
b7h1.cutoff          0     0     0     0     0     0
imp3.cutoff          0     2     2     2     3     3
caix.cutoff          0     0     0     0     0     0
ki67.cutoff          0     0     2     3     2     4
pd1.cutoff           0     0     0     0     0     0
survivin.cutoff      0     0     2     3     2     2
b7h3.cutoff          0     2     3     4     5     6
topo.cutoff          0     0     0     0     0     0
bap1.cutoff          0     0     0     0     0     0
pbrm1.cutoff         0     0     2     2     2     2
SSIGN                0     2     3     4     5     6
age_at_surg          0     2     3     2     5     6
```


b7h3.cutoff > 0.00
SSIGN > 4.00
SSIGN > 5.00
survivin.cutoff > 0.00
age_at_surg > 72.02
pbrm1.cutoff > 0.00
age_at_surg > 66.25
p3: 1.84
ki67.cutoff > 0.00
p3: 1.84
age_at_surg > 57.89
p3: 1.84
p2: 5.19
imp3.cutoff > 0.00
ki67.cutoff > 0.00
p2: 5.19
SSIGN > 6.00
p4: 6.41
p2: 5.19
p2: 5.19
age_at_surg > 72.75
p6: 11.14
p1: 7.18
p3: 1.84
p2: 5.19
age_at_surg > 77.41
p1: 7.18
p2: 5.19
p5: 9.93

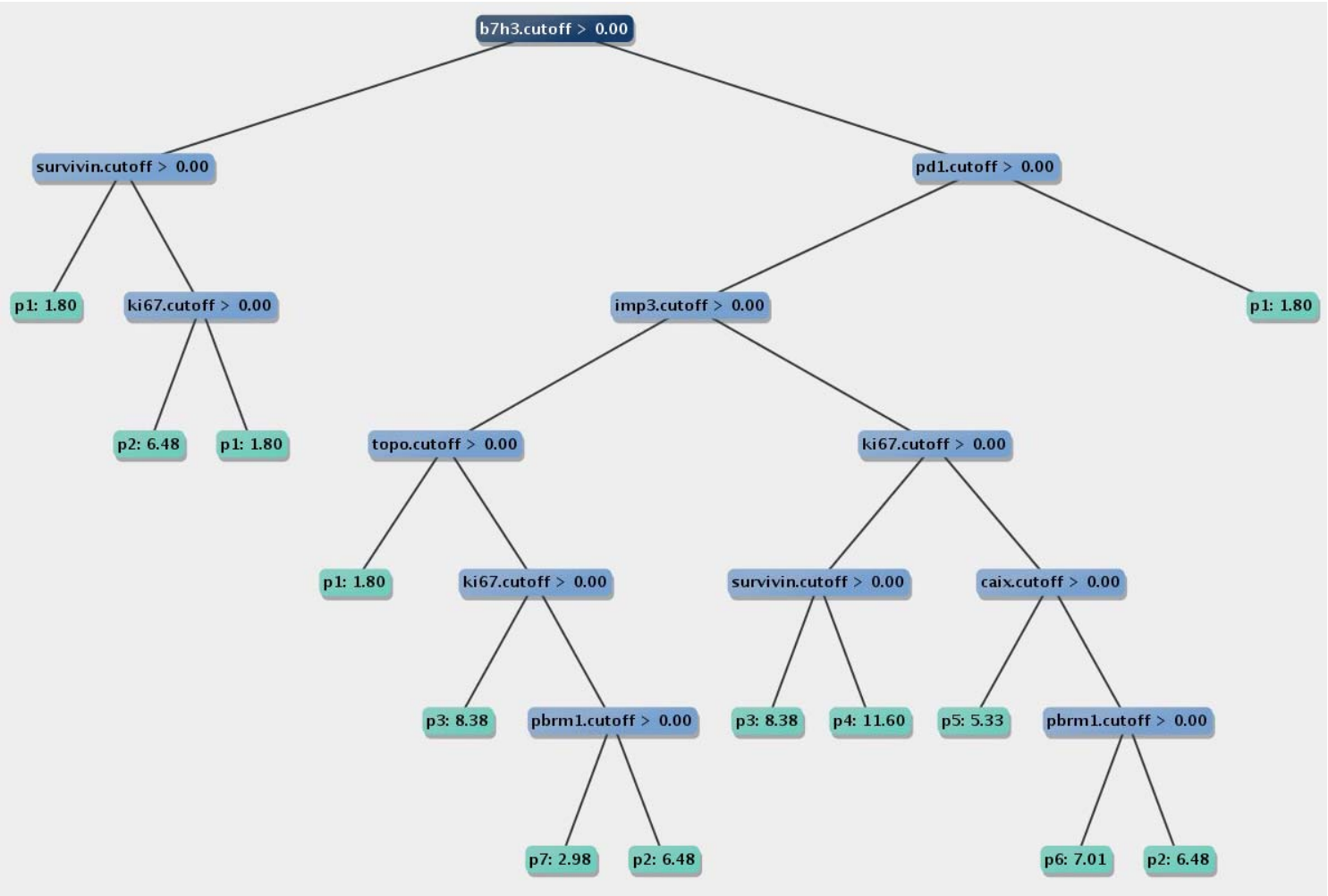
partDSA output with age/SSIGN/10-markers - Brier weighting scheme

```
partDSA object
# partitions   mean CV error   sd CV error   test risk
1              0.652631        0.040464      0.659870
2              0.479782        0.126455      0.415294
3              0.486350        0.140408      0.384646
4              0.474728        0.141065      0.382586
5              0.480399        0.133865      0.382439
6              0.473335        0.116014      0.366077


Variable importance matrix:
                COG=1 COG=2 COG=3 COG=4 COG=5 COG=6
b7h1.cutoff         0     0     0     0     0     0
imp3.cutoff         0     0     0     0     0     0
caix.cutoff         0     0     0     0     0     0
ki67.cutoff         0     0     0     0     0     2
pd1.cutoff          0     0     0     0     0     0
survivin.cutoff     0     0     2     2     3     3
b7h3.cutoff         0     2     2     2     2     2
topo.cutoff         0     0     0     0     0     0
bap1.cutoff         0     0     0     0     0     0
pbrm1.cutoff        0     2     2     2     2     2
age_at_surg         0     2     3     4     4     6
SSIGN               0     2     3     4     5     6
```



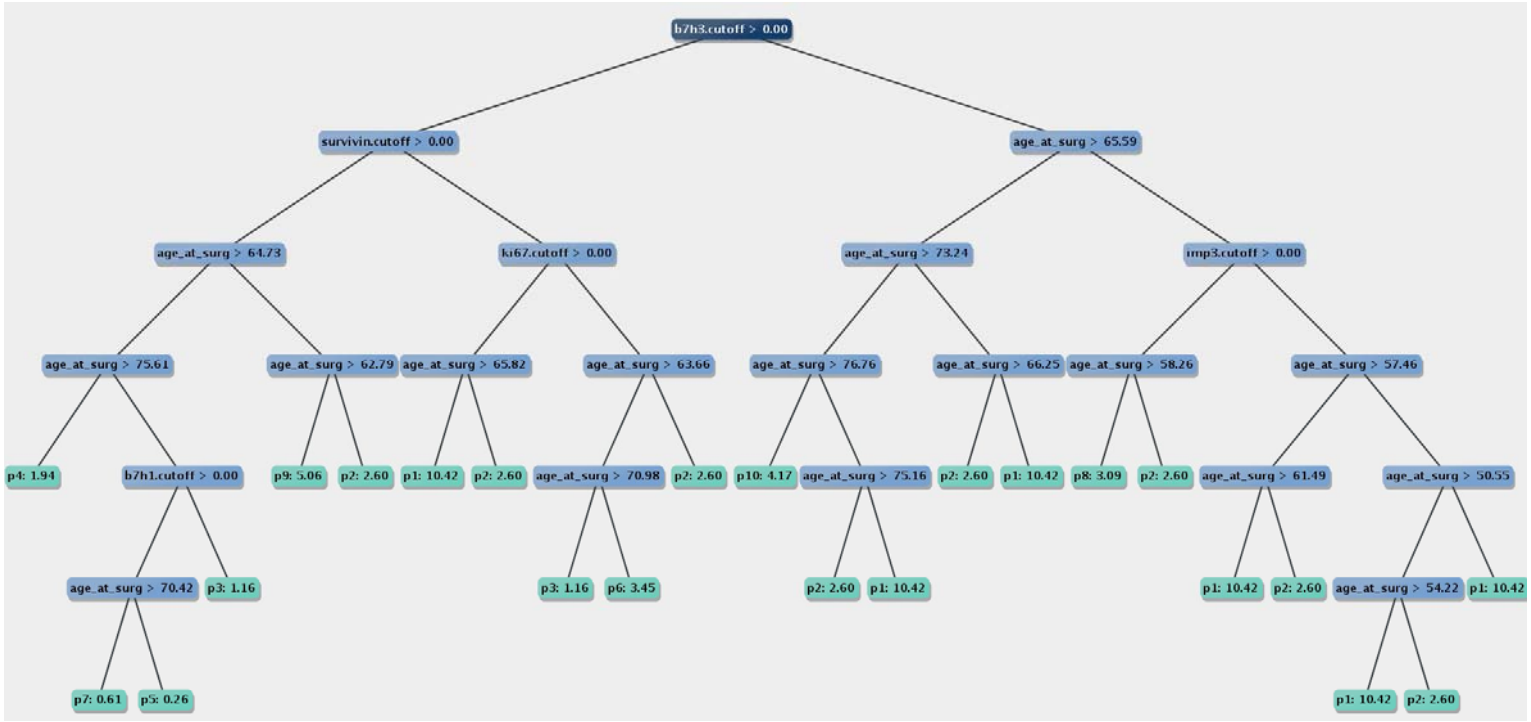
partDSA output with only the 10-markers - IPCW weighting scheme

```
partDSA object
# partitions   mean CV error   sd CV error   test risk
1              23.309061       13.605709     24.933713
2              20.666010       15.610119     15.691585
3              21.243106       14.707665     14.307332
4              20.369243       13.238090     14.006881
5              20.600379       13.561260     13.962062
6              20.850989       13.528225     13.931341
7              20.737297       13.319620     13.912869
8              20.740429       13.135417     14.054967
9              20.668631       13.021742     14.022200
10             20.638699       12.888644     13.998815

Variable importance matrix:
                COG=1 COG=2 COG=3 COG=4 COG=5 COG=6 COG=7 COG=8 COG=9 COG=10
b7h1.cutoff         0     0     0     0     0     0     0     2     3      3
imp3.cutoff         0     2     3     4     5     6     7     5     5      6
caix.cutoff         0     0     0     0     2     2     3     2     4      4
ki67.cutoff         0     2     3     4     5     6     7     5     5      6
pd1.cutoff          0     2     3     4     5     6     7     8     9      9
survivin.cutoff     0     2     3     4     4     4     4     6     7      8
b7h3.cutoff         0     2     3     4     5     6     7     8     9     10
topo.cutoff         0     2     2     3     3     4     4     2     2      2
bap1.cutoff         0     0     0     0     0     0     0     0     0      0
pbrm1.cutoff        0     2     2     2     2     2     3     3     3      3
```



partDSA output with age and the 10-markers - IPCW weighting scheme

```
partDSA object
# partitions   mean CV error   sd CV error   test risk
1              22.772229       12.225680     24.933713
2              25.436115       14.309642     11.395835
3              23.839822       13.365276     10.964599
4              24.022321       13.215157     10.868257
5              23.918442       12.937799     10.850429
6              23.767178       12.992911     10.842091
7              24.476192       13.218271     10.836790
8              24.556313       13.170692     10.835194
9              24.576160       13.245505     10.835188
10             24.040116       12.646550     10.650231
Variable importance matrix:
                COG=1 COG=2 COG=3 COG=4 COG=5 COG=6 COG=7 COG=8 COG=9 COG=10
b7h1.cutoff         0     0     0     0     0     0     0     0     0      3
imp3.cutoff         0     2     2     2     2     2     2     3     3      3
caix.cutoff         0     0     0     0     0     0     0     0     0      0
ki67.cutoff         0     2     3     3     4     4     4     4     4      4
pd1.cutoff          0     0     0     0     0     0     0     0     0      0
survivin.cutoff     0     2     3     3     4     5     5     5     5      8
b7h3.cutoff         0     2     3     4     5     6     7     8     9     10
topo.cutoff         0     0     0     0     0     0     0     0     0      0
bap1.cutoff         0     0     0     0     0     0     0     0     0      0
pbrm1.cutoff        0     0     0     2     2     2     2     2     2      0
age_at_surg         0     2     3     4     5     6     7     8     9     10
```



Construction of BioProg Score

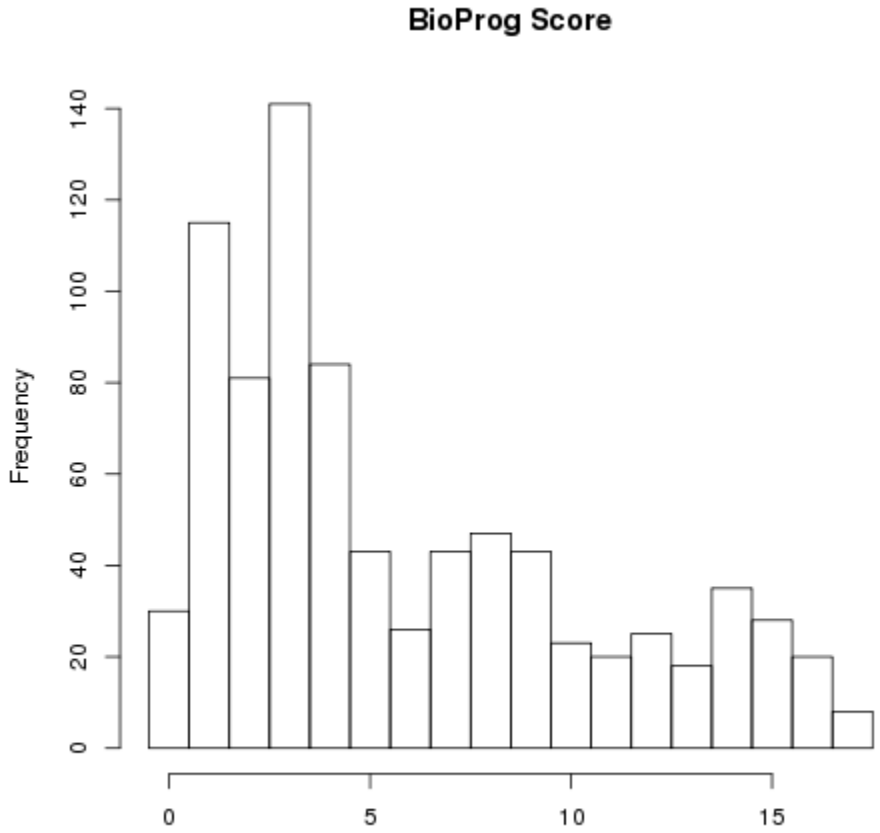
Ultimately, after much deliberation, we used the positively-constrained Ridge Regression coefficients to build our scoring algorithm for BioPROG. Breifly, we start by dividing each coefficient by the highest coefficient, that of IMP3 (0.872). We then multiply the result by 5 and round to the nearest integer. Anyway, as it's a linear shift it shouldn't matter one way or another. Muliplying by five in this case allows the genes with low coefficients to contribute 1 to the overall BioProg score, which is satisfying.

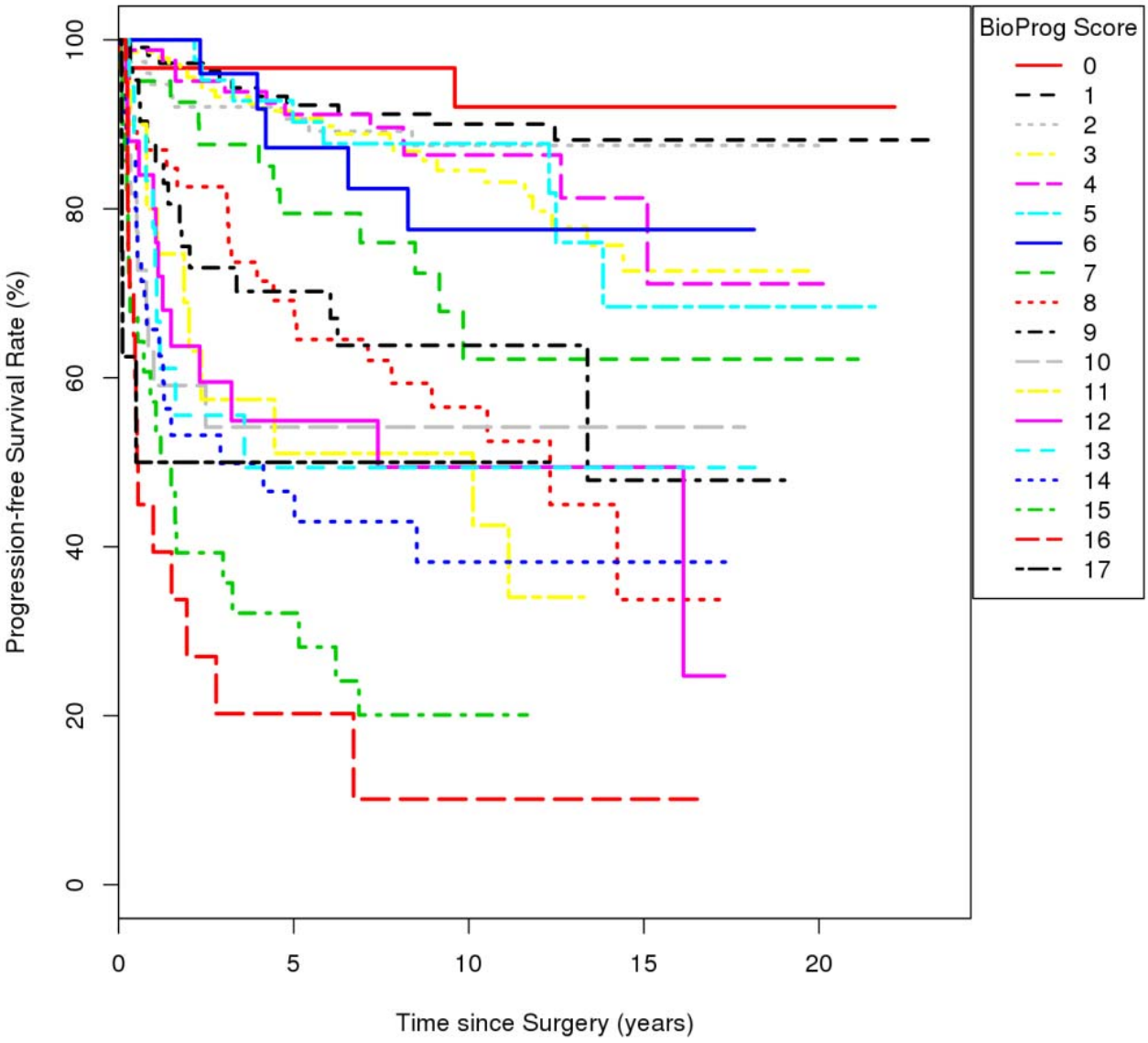
Table of numbers from each step below.

gene	coef	div	mult	score
b7h1	0.4123306	0.4728466	2.3642329	2
imp3	0.8720177	1.0000000	5.0000000	5
caix	0.2390133	0.2740922	1.3704611	1
ki67	0.1441463	0.1653021	0.8265104	1
pd1	0.2123010	0.2434595	1.2172975	1
survivin	0.6577184	0.7542489	3.7712444	4
b7h3	0.1337130	0.1533375	0.7666875	1
topo	0.1218399	0.1397218	0.6986090	1
bap1	0.0000000	0.0000000	0.0000000	0
pbrm1	0.3089554	0.3542995	1.7714973	2

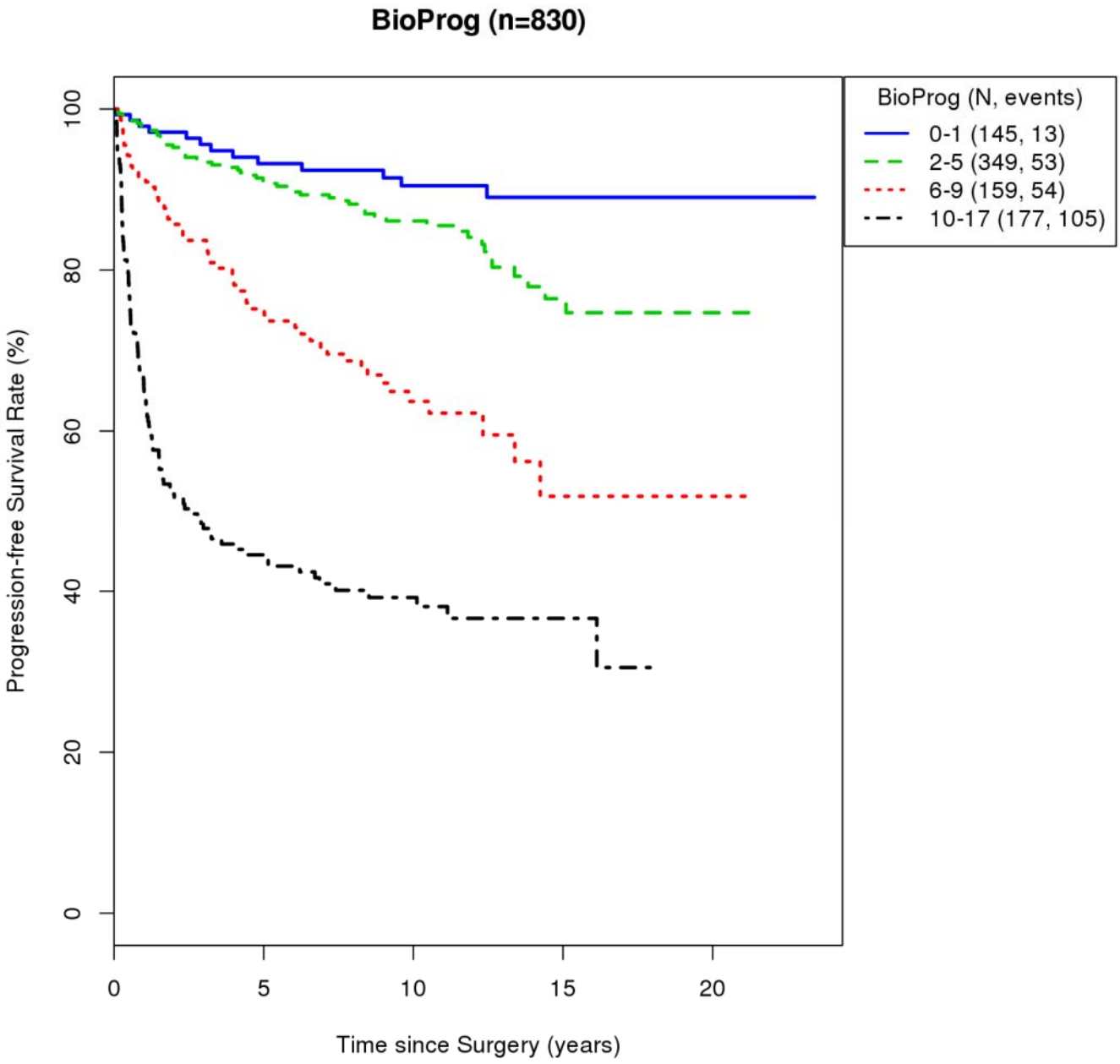
BioProg Score

Score	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
N	30	115	81	141	84	43	26	43	47	43	23	20	25	18	35	28	20	8
Events	2	11	9	24	12	8	5	12	22	15	10	11	13	9	20	22	16	4

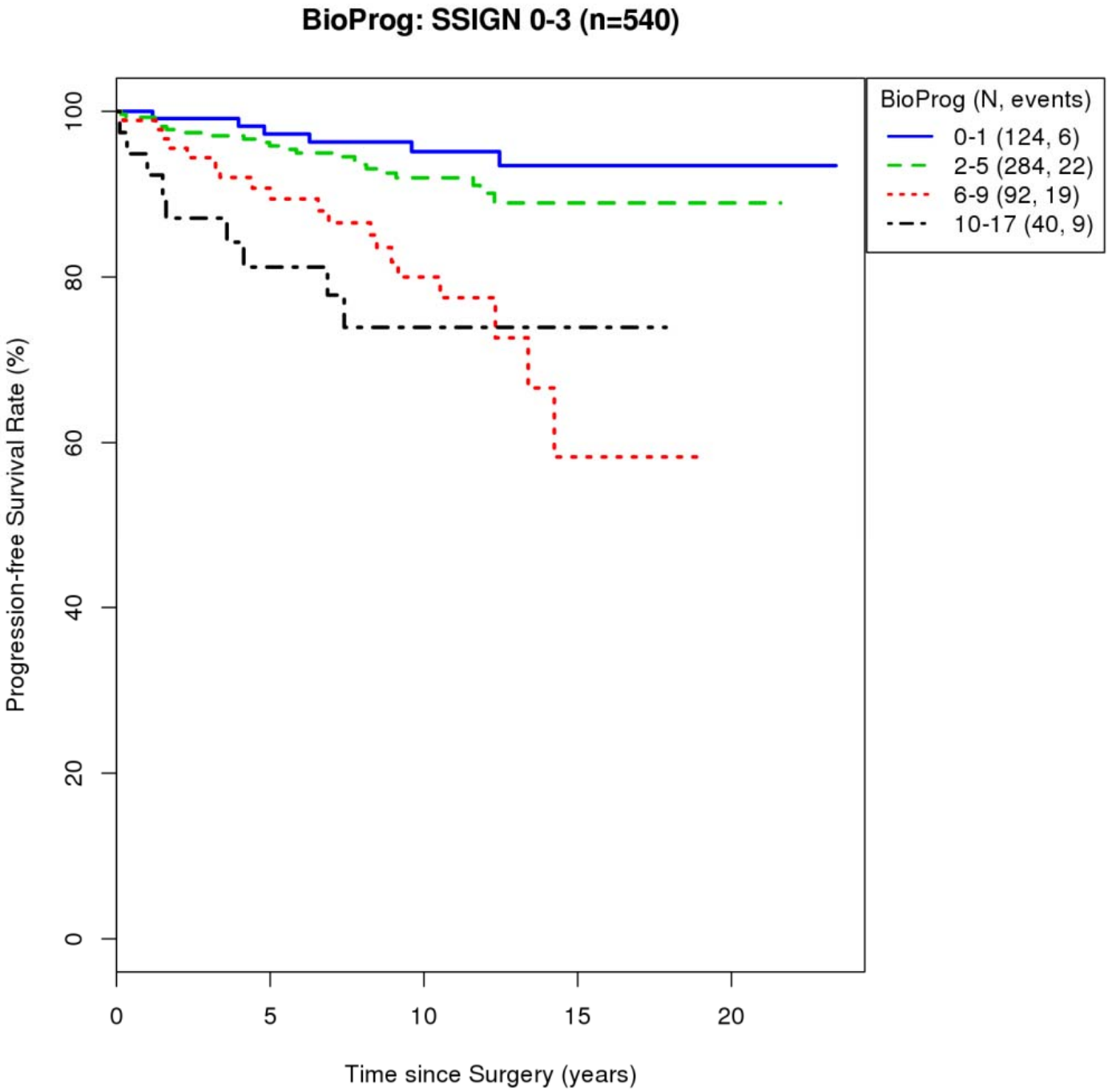



BioProg (n=830)
Progression-free Survival Rate (%)
0
20
40
60
80
100
0
5
10
15
20
Time since Surgery (years)
BioProg Score
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17

Whiel the above Kaplan Meier cureve show the survivla expereince for all 17 individual possible scores of BioPROG, to promote true clincally utility, we group BioProg in to 4 distinct categories, shown here:



Here we demonstrate the ability of BioPROG to further refine prognosis even among the very homogneous group of low SSIGN score patients (i.e those with a SSIGN score of 0-3). Recall, the SSIGN score is a clicnal algorithm that can place patients in to low, intermediate and high risk categories for RCC death. Here, we demostrate that our new BioPROG core an provide further refinement of that prognosis even among those detmeine to be a"low risk" based ont eh SSIGN score.



BioProg Cox Models Showing the independent association of our Four BioPROG categories with progression-free survival overall and among those with low SSIGN scores.

		Progression-free Survival (n=830, 225 events)			PFS in SSIGN 0-3 (n=540, 56 events)		
		HR (95% CI)	P-value	c-index	HR (95% CI)	P-value	c-index
Adjusted for Age	Continuous BioProg	1.21 (1.18 - 1.24)	<2.0E-16	0.758	1.17 (1.10 - 1.25)	2.64E-07	0.686
Adjusted for Age and SSIGN	Continuous BioProg	1.07 (1.04 - 1.11)	4.73E-05	0.845	NA	NA	
Adjusted for Age	BioProg 0-1	1.0 (reference)	NA	0.750	1.0 (reference)	NA	0.677
	BioProg 2-5	1.86 (1.01 - 3.43)	0.045		1.73 (0.69 - 4.32)	0.24	
	BioProg 6-9	4.86 (2.64 - 8.95)	3.76E-07		5.18 (2.05 - 13.12)	0.00052	
	BioProg 10-17	12.32 (6.86 - 22.11)	<2.0E-16		6.46 (2.25 - 18.56)	0.00053	
Adjusted for Age and SSIGN	BioProg 0-1	1.0 (reference)	NA	0.844	NA	NA	
	BioProg 2-5	1.56 (0.85 - 2.86)	0.16		NA	NA	
	BioProg 6-9	2.69 (1.45 - 5.00)	0.0017		NA	NA	
	BioProg 10-17	3.08 (1.64 - 5.78)	0.00048		NA	NA	

Final BioProg Score - excluding genes that are non-zero < 90% bootstraps

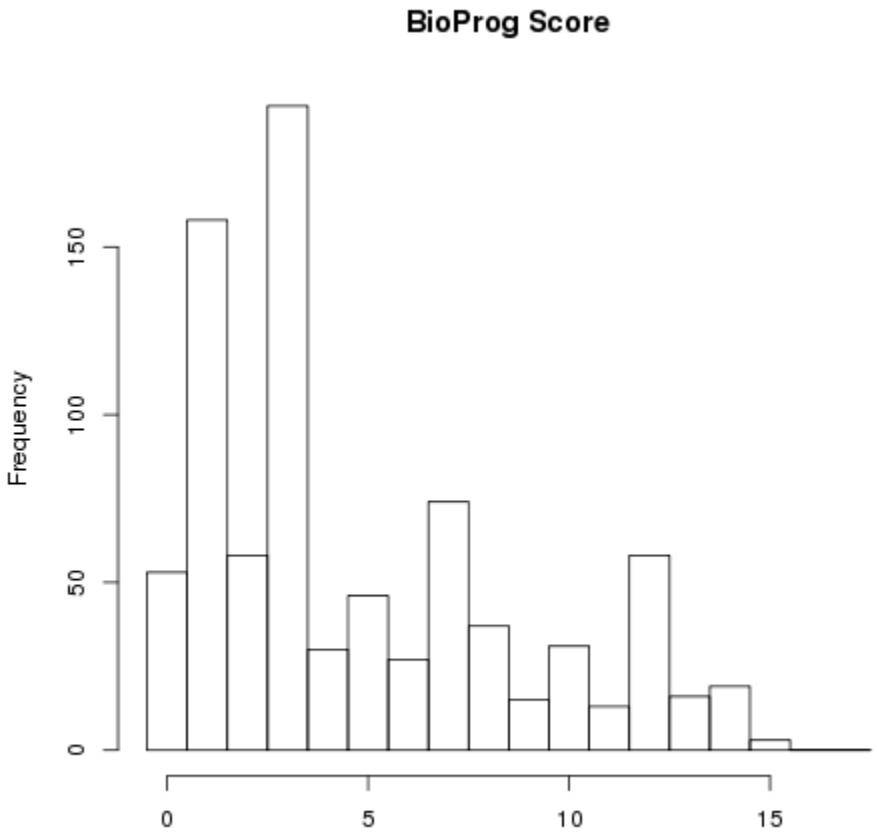
To determine the most cost-effective (and therefore likely to be used clinically) BioPROG model, we evaluated a more parsimonious BioProg score where the genes that are non-zero in bootstraps < 90% of the time are excluded from the score (we set ki67, b7h3, and topo2 to multiplier=0; bap1 already was excluded).

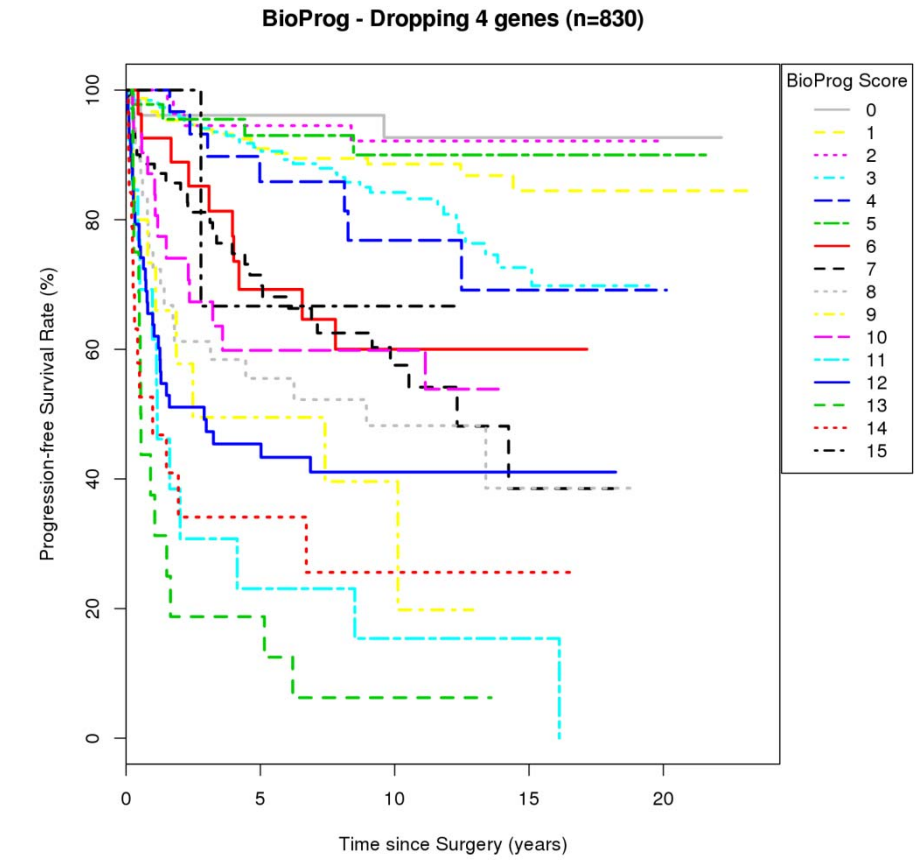
A table of multipliers is below.

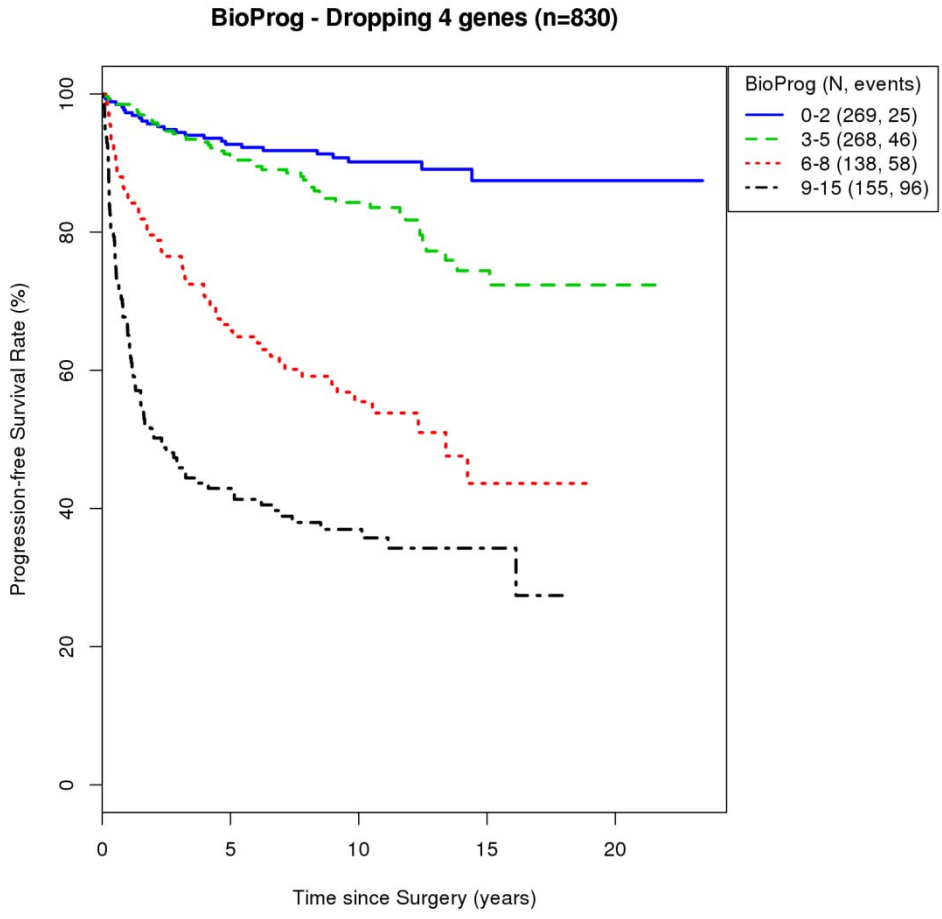
gene	coef	div	mult	score
b7h1	0.4123306	0.4728466	2.3642329	2
imp3	0.8720177	1.0000000	5.0000000	5
caix	0.2390133	0.2740922	1.3704611	1
ki67	0.1441463	0.1653021	0.8265104	0
pd1	0.2123010	0.2434595	1.2172975	1
survivin	0.6577184	0.7542489	3.7712444	4
b7h3	0.1337130	0.1533375	0.7666875	0
topo	0.1218399	0.1397218	0.6986090	0
bap1	0.0000000	0.0000000	0.0000000	0
pbrm1	0.3089554	0.3542995	1.7714973	2

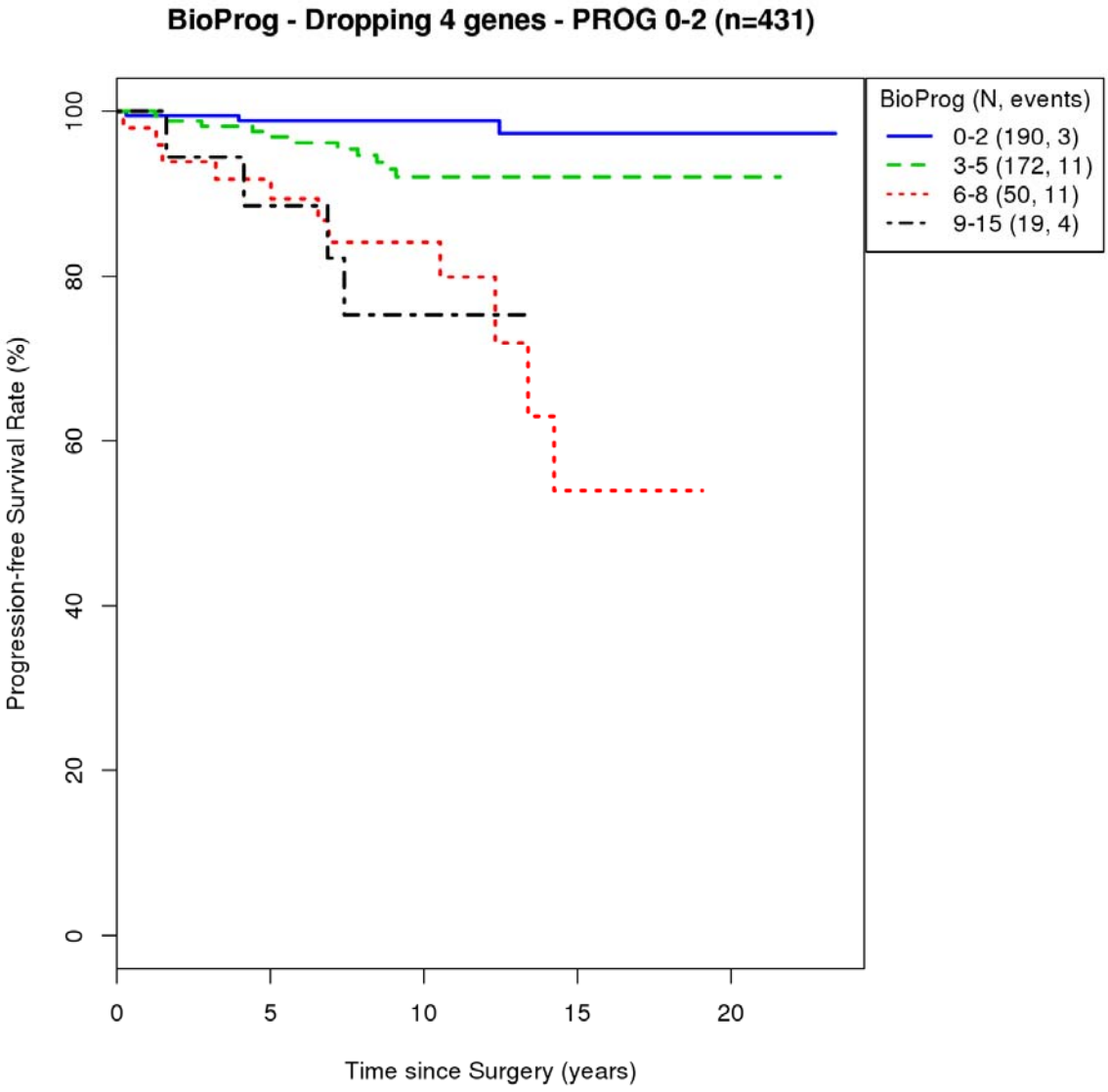
BioProg Score

Score	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
N	53	158	58	192	30	46	27	74	37	15	31	13	58	16	19	3
Events	3	18	4	35	7	4	10	29	19	9	13	12	33	15	13	1




BioProg - Dropping 4 genes (n=830)
BioProg Score
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Progression-free Survival Rate (%)
0
20
40
60
80
100
0
5
10
15
20
Time since Surgery (years)


BioProg - Dropping 4 genes (n=830)
BioProg (N, events)
0-2 (269, 25)
3-5 (268, 46)
6-8 (138, 58)
9-15 (155, 96)
Progression-free Survival Rate (%)
0
20
40
60
80
100
0
5
10
15
20
Time since Surgery (years)


BioProg - Dropping 4 genes - PROG 0-2 (n=431)
BioProg (N, events)
0-2 (190, 3)
3-5 (172, 11)
6-8 (50, 11)
9-15 (19, 4)
Progression-free Survival Rate (%)
0
20
40
60
80
100
0
5
10
15
20
Time since Surgery (years)

Final BioProg Cox Models - after dropping 4 biomarkers

		Progression-free Survival (n=830, 225 events)			PFS in PROG 0-2 (n=431, 29 events)		
		HR (95% CI)	P-value	c-index	HR (95% CI)	P-value	c-index
Adjusted for Age	Continuous BioProg	1.24 (1.21 - 1.28)	<2.0E-16	0.764	1.30 (1.17 - 1.44)	6.66E-07	0.750
Adjusted for Age and PROG	Continuous BioProg	1.08 (1.04 - 1.13)	1.94E-05	0.852	NA	NA	
Adjusted for Age	BioProg 0-2	1.0 (reference)	NA	0.756	1.0 (reference)	NA	0.757
	BioProg 3-5	1.88 (1.15 - 3.06)	0.012		3.99 (1.11 - 14.39)	0.034	
	BioProg 6-8	5.88 (3.67 - 9.44)	1.92E-13		15.27 (4.22 - 55.28)	3.27E-05	
	BioProg 9-15	11.77 (7.53 - 18.39)	<2.0E-16		15.82 (3.46 - 72.24)	3.67E-04	
Adjusted for Age and PROG	BioProg 0-2	1.0 (reference)	NA	0.853	NA	NA	
	BioProg 3-5	1.42 (0.87 - 2.33)	0.16		NA	NA	
	BioProg 6-8	2.94 (1.81 - 4.78)	1.31E-05		NA	NA	
	BioProg 9-15	2.92 (1.77 - 4.82)	2.70E-05		NA	NA	

Conclusion: From our efforts in this task, we have develop a parsimonious collection of biomarkers that when their expression in turmo tissue is combined together, produce a scoring algorithm (BioPROG) that separates ccRCC patients into distinct groups based on their survival probability. Most notably, BioPROG has the ability to parse out survicla experience even among those who have alrady been determined to be at "low risk" of RCC death based on the widely used Mayo SSIGN score (i.e. SSIGN score of <3). This panel of biomarkes is currenly being developed in our Department of Laboratory Medicine and Pathology as a clinically available CLIA-approved test.

Task 4. **Identify differentially expressed proteomic biomarkers by MALDI mass spectrometry imaging in primary ccRCC tumor tissues matched to ccRCC tissue metastatic to the lung from the same person.**

a. Perform MALDI-TOF MS imaging analysis of 15 pairs of primary and metastatic renal tissues (30 total)
b. Analyze peak intensity data inter- and intra-sample to determine differentially expressed proteins and peptides

MILESTONE: Establish a panel of 5-10 differentially expressed peak markers that distinguish the primary and metastatic tumor pairs. Months 20-36 (Drake) To be done at MUSC

Experimental Progress: Dr. Parker has been trying to obtain these tissues from the Mayo Clinic Rochester, with little success. At the time of submission, these samples were available. Subsequent issues related to consent and permitted use of this specific sample cohort have precluded getting them for the proposed analyses. In the interim, we have begun analysis of 20 RCC tissues from recurrent and non-recurrent tumors, but tissues that contain extensive regions of tumor, margin areas and adjacent non-tumor regions. These tissues were provided by longtime collaborator Dean Troyer, M.D., a GU pathologist at EVMS. Matched primary and metastatic tumor pairs are also being obtained from EVMS, from Dr. Troyer and the urologic surgeon there, Ray Lance, M.D. An emphasis on lipid, protein and glycan analysis has been done.

Our preliminary studies have identified multiple glycan and glycoprotein species uniquely expressed in the tumor margin regions, but not expressed in the non-tumor or tumor regions of the same tissue. This is the result of using a novel method to profile N-linked glycans directly on tissue using MALDI-MS imaging following PNGaseF digests. Depending on the tissue, 30-40 N-glycan species can be simultaneously detected by this method. Individual glycan and/or lipid ion intensities are converted to a color pixel scale for creating an image, linked directly to the histopathology of the tissue. Using the unique tumor/margin RCC tissues, four patterns of N-glycan and lipid specie expression have been observed, those present in: 1) the immediate margin area of non-tumor tissue adjacent to tumor; 2) only in non-tumor regions; 3) only in tumor regions; or 4) primarily in tumor regions but extended beyond the margin. An initial analysis of glycoproteins present in these three regions has also been done using new HCD-PD-ETD glycopeptide sequencing workflows. We hypothesize that optimization of this experimental workflow can be used to identify specific glycoprotein and glycan biomarkers indicative of the many changes that occur during the transition of organ confined tumors to the metastatic phenotype. Target molecules, including proteins, will be further analyzed in the primary tumor and metastatic disease pairs. Examples of this data are provided in the Appendix section for Figures 1-8.

Task 5. Identify differentially expressed metabolite biomarkers in primary ccRCC tumor tissues matched to ccRCC tissue metastatic to the lung from the same person.

a. Perform metabolite analysis of 15 pairs of primary and metastatic renal tissues (30 total), to be done by Metabolon, Inc.
b. Determine metabolite differences between primary and metastatic tissues.

MILESTONE: Identify a panel of metabolites from tissue that distinguish primary and metastatic renal tumor pairs. Months 12-18 (Metabolon, Parker)

As discussed above, we successfully completed the metabolic analysis of renal tissue samples (and now plasma samples as well) obtained from the Mayo Clinic. Below we highlight the key results from these analyses:

• **_Tumor vs. normal tissue comparison.** A comparison of tumor and normal tissue metabolites is valuable to help understand the direction of changes which occur as tumor cells reprogram their metabolic pathways to support proliferation. Consistent with previous observations for a broad range of tumor cells, the RCC cells appeared to have shifted metabolic pathways to support rapid ATP generation through enhanced glycolysis, while increasing biosynthesis and managing cellular redox status. Increased levels of **glucose-6-phosphate (G-6-P)** and **fructose-6-phosphate (F-6-P)** are representative of an increase uptake of **glucose** to support enhanced glycolysis. Some of the glucose can be utilized to support greater flow through the Pentose Phosphate pathway (PPP), which generates NADPH for biosynthesis and antioxidant protection against reactive oxygen species (ROS). This increased PPP activity was evidenced by increases in **ribose, ribose-5-phosphate and gluconate**. Interestingly, although **G-6-P** and **F-6-P** were increased, some downstream metabolites such as **3-phosphoglycerate 2-phosphoglycerate** and **phosphoenolpyruvate** were observed to be lower in tumor, relative to normal kidney tissue. This decrease in 3-carbon glycolytic intermediates may be linked to the low **NAD**+ levels in the tumor tissue.

NAD is required by glyceraldehyde-3-phosphate dehydrogenase to convert **glyceraldehyde-3-phosphate** to **1-3-bisphosphoglycerate** in glycolysis. The decline in **NAD**+ levels may be reflective of hypoxic conditions in the tumor which are known to decrease the NAD+/NADH ratio. **NAD**+ can also be consumed by poly-ADP-ribose polymerase in DNA repair. An alternative explanation for lowered downstream glycolysis metabolites is elevated **citrate,** which was observed in the tumor tissues. Citrate along with ATP is a negative regulator of phosphofructokinase-1 which would decrease conversion of **fructose-6-phosphate** to **fructose 1,6-bisphosphate**. Thus, this potential decrease in phosphofructokinase-1 activity could result in a buildup of fructose-6-phosphate and be a sign of high levels of ATP and/or **citrate**. Combined, low **NAD**+ and high **citrate** may be signs of an energy surplus. Amino acids are in general decreased in tumor versus normal tissue and this is consistent with greater demand for protein synthesis.
Elevated citrate, which can serve as a fatty acid precursor compound, and an accompanying increase in fatty acid biosynthetic pathways frequently are components of a tumor cell metabolic signature. In this study, elevation in citrate was accompanied by an increase in fatty acids, likely marking increased fatty acid biosynthesis as would be required to support cell proliferation. In addition, there was a decline in **lysolipids** in tumor compared to non-tumor tissue, which supports reduced fatty acid catabolism for energy generation and decreased membrane turnover.
Tumor cell metabolism also is known frequently to show alterations in pathways important for redox homeostasis. Consistent with known tumor cell metabolism alteration, was the increase in

glutathione, both oxidized and reduced (in comparing all tumor vs. all normal) observed in this study. This elevation in the level of a key cellular anti-oxidant typically reflects the higher anti-oxidant demand to quench reactive oxygen species produced during rapid cell proliferation.

N-acetyl amino acids are strongly decreased in progressing tumors. One metabolite class that was observed to associate with tumor progression is a decrease in N-acetyl amino acids. This is most dramatic for **N-acetyl glutamine** (p=0.047 Int-prog vs. Int-nonprog) and **N-acetyl asparagine** (p=0.041 Int-prog vs. Int-nonprog). Curiously, these two metabolites were increased in tumor relative to normal tissue, yet differentiated progressive vs. non-progressive by being decreased in progressing tumors. N-acetyl amino acids are salvaged in the kidney through conversion to their parent amino acid by the enzyme aminoacylase-1. This enzyme is expressed most abundantly expressed in the kidney tubular epithelium. One might hypothesize that an upregulation of aminoacylase-1 activity could serve to increase the efficiency of amino acid salvage, for energetic or biosynthetic purposes. The upregulation of aminoaylase-1 activity would presumably decrease the level of the N-acetylated amino acids present in tumor cells with this increased activity. Alternatively, the decrease in N-acetyl amino acids may represent a change in transport efficiency.

• **A fibrinogen cleavage peptide is more abundant in progressing tumors.** One of the interesting metabolites significantly increased in progressing tumors was a fibrinogen cleavage peptide **ADSGEGDFXAEGGGVR**. Its level was up 5.7X comparing intermediate progressing vs. intermediate non-progressing tumor (p-value = 0.008). It was also increased >4X in high progressing vs. intermediate non-progressing (p-value = 0.02). It is possible that this change in cleavage peptide levels is associated with remodeling or breakdown in the extra cellular matrix of the tumor stroma. One might postulate that more aggressive tumors would secrete more proteases, breaking down fibrinogen and permitting expansion of the tumor cell population along with better access to the vasculature. Tests for fibrinogen breakdown products are already utilized in clinical oncology. It is also possible that increased fibrinogen breakdown peptide is reflective of alterations in clotting homeostasis, within the discontinuous leaky vasculature frequently observed in tumors.

• **The histamine metabolite 1-methylimidazoleacetate was decreased in progressing tumors.** While not a robust differentiating metabolite, **1-methylimidazoleacetate** is of interest because it does show a decline in progressing tumors and there are published results that histamine may restrict cancer growth. One can postulate that if this metabolite represents decreased histamine in the tumor, then the decline in **histamine** might be linked to tumor aggressiveness. **Histamine** was also measured in some samples, but was highly variable and frequently not detected at all. One can categorize this metabolite's link to tumor progression as interesting, but highly speculative.

Comparison of metabolites across SSIGN groups reveals a gradation of changes associated with tumor metabolism changes. In comparing metabolites that define the SSIGN groupings, little stood out as a step function or unique to any grade. The metabolite changes that were observed between tumor and normal tissue were graduated across the SSIGN scores, with the level of change being more pronounced as the SSIGN score increased. Some individual metabolites did reach statistical significance when comparing low score vs. intermediate score, but these were often metabolites with variable detection across samples in the groups (low % fill).

Several unnamed biochemicals were correlated with progression. Among the unnamed biochemicals, a number were correlated with progression at a statistically significant level. X11381 was one of the unknowns that declined comparing inter-progressing vs. inter-nonprogressing (↓ 54%;

p-value = 0.0043) and high-progressing vs. inter-nonprogressing (↓66%; p-value = 0.0019). Two unnamed biochemicals increased in progressing vs. non-progressing. X-16103 increased 25%

and 34% (p-values: 0.023 and 0.005) for inter-prog vs. inter-nonprog and high-prog vs. inter-nonprog, respectively. X-16209 increased to similar extents and with similar p-values to X-16103. Thus, there were examples of increasing and decreasing unknowns which tracked with progression.

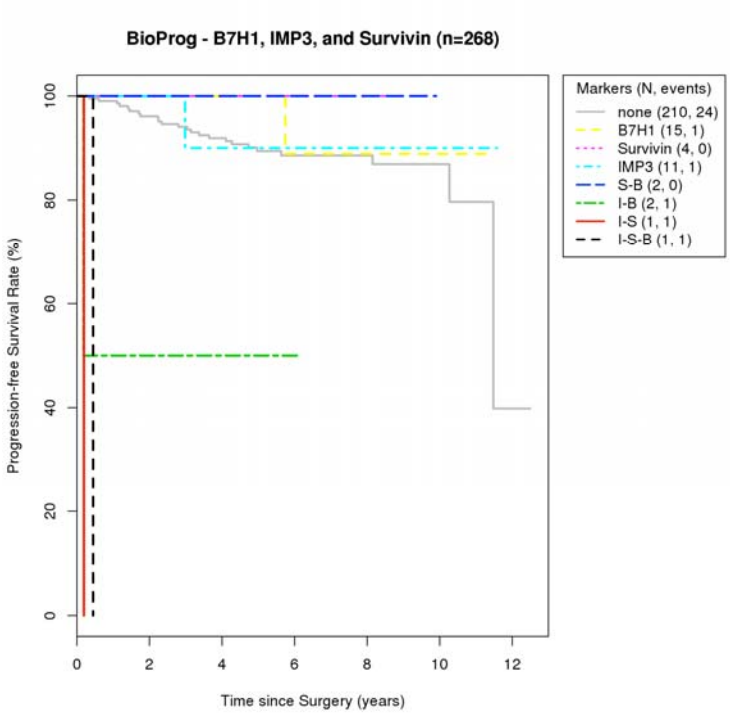
As discussed in our grant application, we will now seek to validate these tissue-based metabolites and determine if they can be utilized as a valuable piece of a multivariable biomarker algorithm to help predict RCC recurrence following surgery. Of note, we were able to identify **several plasma-based biomarkers** that are also correlated with RCC aggressiveness. Since this work was not part of the original project, we do not include these results but only mention again that they will complement the above analyses and push us further down the road toward a clinically valuable tool for RCC

Task 6. **Independently validate the differential expression of the candidate biomarkers identified in Tasks 3, 4 and 5 and estimate the association of their expression in metastatic ccRCC tissue with time to death.**

a. Examine the co-expression of the subset of existing seven-member biomarker panel (*survivin, B7-H1, B7-H4, Ki-67, IGF-IR, IMP-3 and CA-IX*) incorporated into the BioMarker prediction algorithm in a cohort of over 500 samples.
b. Assess the collinearity between all biomarkers using multivariable Cox proportional hazards regression to evaluate if the full panel of biomarkers are independent predictors of time to death.
c. Evaluate whether MALDI-MS tissue biomarker and metabolite candidates from Tasks 4 and 5 are linked to time of death and inform on molecular pathways for therapeutic intervention.

BioPROG validation

We have validated each of our individual biomarkers but more importantly also our final BioPROG scoring algorithm. Indeed, is was his validation that allowed us to move forward with the development of the clinical version of this test in our Department of Laboratory Medicine (external validation is required to move forward. Below is the KM curve from this key validation effort.



Primary vs Metastatic for Key Biomarkers: PBRM1, BAP1, Ki-67, and TopoII

We assessed tissue for 187 metastases from 114 patients (60 M0 and 54 M1). Of these, 177 mets were stained for BAP1 and PBRM1, and the other 10 (7 M0 and 3 M1) had no viable tumor. We also sent the 52 available M1 patients for staining, and all of these were successful (though a few didn't stain well, an assessment was still made).

Below is a table describing the available slides on these 114 patients that we analyzed.

Available Slides

	M0	M1	Overall
Primary Tumors	60	54	114
Number of Metastases			
1	32	39	71
2	16	12	28
3	5	2	7
4	6	0	6
7	1	0	1
8	0	1	1
Total	110	77	187

Below a table of usable slides by each biomarker we evaluated.

PBRM1: those with a successful Primary and Met

	M0	M1	Overall
Primary Tumors	48	51	99
Number of Stained Mets*			
1	24	38	62
2	15	10	25
3	5	2	7
4	3	0	3
6	1	0	1
7	0	1	1
Total	87	71	158

BAP1: those with a successful Primary and Met

	M0	M1	Overall
Primary Tumors	49	50	99
Number of Stained Mets*			
1	25	37	62
2	15	10	25
3	5	2	7
4	3	0	3
6	1	0	1
7	0	1	1

Total	88	70	158

TOPOII: those with a successful Primary and Met

	M0	M1	Overall
Primary Tumors	42	52	94
Number of Stained Mets*			
1	20	38	58
2	13	11	24
3	3	2	5
4	5	0	5
7	1	1	2
Total	82	73	155

Ki-67: those with a successful Primary and Met

	M0	M1	Overall
Primary Tumors	59	54	113
Number of Stained Mets*			
1	33	40	73
2	16	11	27
3	5	2	7
4	4	0	4
7	1	1	2
Total	103	75	178

*i.e. a patient with 4 mets total may only have 1 met successfully stained for a given gene; this is the difference between these tables and those on page 1

Concordance in PBRM1

Overall (n=158)
Concordance=86.7%

PBRM1		Met	
		Negative	Positive
Primary	Negative	80	9
	Positive	12	57

Removing weak positives and focal negatives (n=140)
Concordance=89.3%

PBRM1		Met	
		Negative	Positive
Primary	Negative	70	7
	Positive	8	55

M0 (n=87)
Concordance=90.8%

PBRM1		Met	
		Negative	Positive
Primary	Negative	48	1
	Positive	7	31

M1 (n=71)
Concordance=81.7%

PBRM1		Met	
		Negative	Positive
Primary	Negative	32	8
	Positive	5	26

Amongst Patients with only 1 Met (n=60)
Concordance=93.3%

PBRM1		Met	
		Negative	Positive
Primary	Negative	33	2
	Positive	2	23

PBRM1 Concordance by Site (n=158)

	Discord	Concord
10=CONTRALATERAL ADRENAL	2	4
11=IPSILATERAL ADRENAL	0	8
12=PANCREAS	2	4
17=THYROID	0	1
18=SKIN	0	4
19=BOWEL	0	1
1=NON-REGIONAL NODES	0	8
20=SPLEEN	0	2
21=MUSCLE	0	1
22=OMENTUM	0	1
2=PULMONARY	13	55
3=LIVER	0	11
4=BONE	0	18
5=HEART	1	0
6=BRAIN	2	8
9=OTHER	1	11

Lung Mets: Concordance=80.9% (88.1% in M0's, 69.2% in M1's)

PBRM1		All Lung Mets (n=68)		M0 Only (n=42)		M1 Only (n=26)	
		Negative	Positive	Negative	Positive	Negative	Positive
Primary	Negative	29	8	22	1	7	7
	Positive	5	26	4	15	1	11

PBRM1 Concordance by Site, removing weak positives and focal negatives (n=140)

	Discord	Concord
10=CONTRALATERAL ADRENAL	1	4
11=IPSILATERAL ADRENAL	0	6
12=PANCREAS	0	4
17=THYROID	0	1
18=SKIN	0	4
19=BOWEL	0	1
1=NON-REGIONAL NODES	0	7
20=SPLEEN	0	2
21=MUSCLE	0	1
22=OMENTUM	0	1
2=PULMONARY	12	48
3=LIVER	0	11
4=BONE	0	16
6=BRAIN	1	8
9=OTHER	1	11

Lung Mets, removing weak positives and focal negatives
Concordance=80.0% (88.2% in M0's, 69.2% in M1's)

PBRM1		All Lung Mets (n=60)		M0 Only (n=34)		M1 Only (n=26)	
		Negative	Positive	Negative	Positive	Negative	Positive
Primary	Negative	22	7	15	0	7	7
	Positive	5	26	4	15	1	11

Discordant lung mets in M0's are only observed going from positive to negative (i.e. the met gained the mutation and lost the protein); discordant lung mets in M1's are mostly negative to positive (i.e. met lacks the mutation of the primary).

PBRM1 Difference Amongst Mets

Of the 37 patients with at least 2 stained mets, only two had mets with differing PBRM1 status. Both those patients' mets are below.

clinic	timing	Site	Metastasectomy.Date	PBRM1.Pri	PBRM1.Met
5023303	M0	2=PULMONARY	5/30/1996	p	n
5023303	M0	2=PULMONARY	12/3/1996	p	n
5023303	M0	2=PULMONARY	4/21/1998	p	n
5023303	M0	2=PULMONARY	4/21/1998	p	p
6173686	M1	2=PULMONARY	3/8/2004	p	n
6173686	M1	2=PULMONARY	3/8/2004	p	p

Concordance in BAP1

Overall (n=158)
Concordance=98.7%

BAP1		Met	
		Negative	Positive
Primary	Negative	40	0
	Positive	2	116

Removing weak positives and focal negatives (n=123)
Concordance=99.2%

BAP1		Met	
		Negative	Positive
Primary	Negative	30	0
	Positive	1	92

M0 (n=88)
Concordance=100.0%

BAP1		Met	
		Negative	Positive
Primary	Negative	23	0
	Positive	0	65

M1 (n=70)
Concordance=97.1%

BAP1		Met	
		Negative	Positive
Primary	Negative	17	0
	Positive	2	51

Amongst Patients with only 1 Met (n=60)
Concordance=98.3%

BAP1	Met	
	Negative	Positive

Primary	Negative	6	0
	Positive	1	53

BAP1 Concordance by Site (n=158)

	Discord	Concord
10=CONTRALATERAL ADRENAL	0	6
11=IPSILATERAL ADRENAL	0	8
12=PANCREAS	0	6
17=THYROID	0	1
18=SKIN	0	4
19=BOWEL	0	1
1=NON-REGIONAL NODES	0	8
20=SPLEEN	0	2
21=MUSCLE	0	1
22=OMENTUM	0	1
2=PULMONARY	1	68
3=LIVER	0	11
4=BONE	1	17
5=HEART	0	1
6=BRAIN	0	9
9=OTHER	0	12

BAP1 Concordance by Site, removing weak positives and focal negatives (n=123)

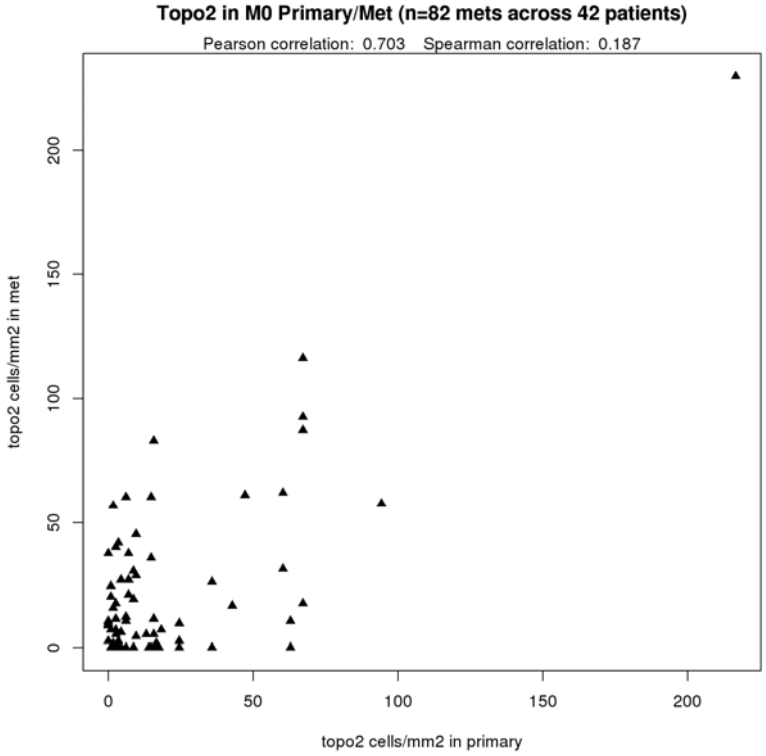
	Discord	Concord
10=CONTRALATERAL ADRENAL	0	5
11=IPSILATERAL ADRENAL	0	6
12=PANCREAS	0	3
17=THYROID	0	1
18=SKIN	0	4
19=BOWEL	0	1
1=NON-REGIONAL NODES	0	7
20=SPLEEN	0	2
21=MUSCLE	0	1
22=OMENTUM	0	1
2=PULMONARY	0	49
3=LIVER	0	7
4=BONE	1	16
5=HEART	0	1
6=BRAIN	0	7
9=OTHER	0	11

BAP1 Difference Amongst Mets

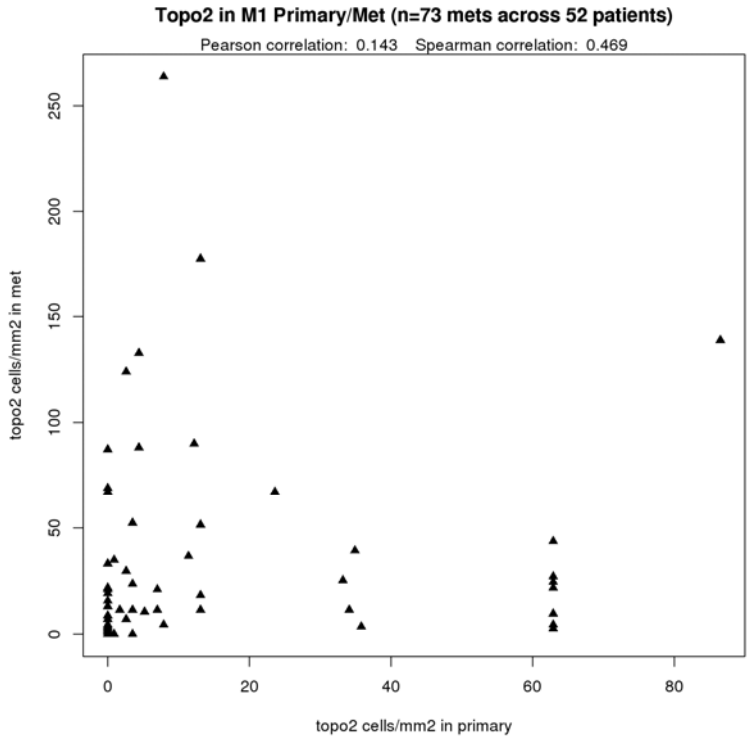
Of the 37 patients with at least 2 mets stained, only one had mets with differing BAP1 status. That patients' mets are below.

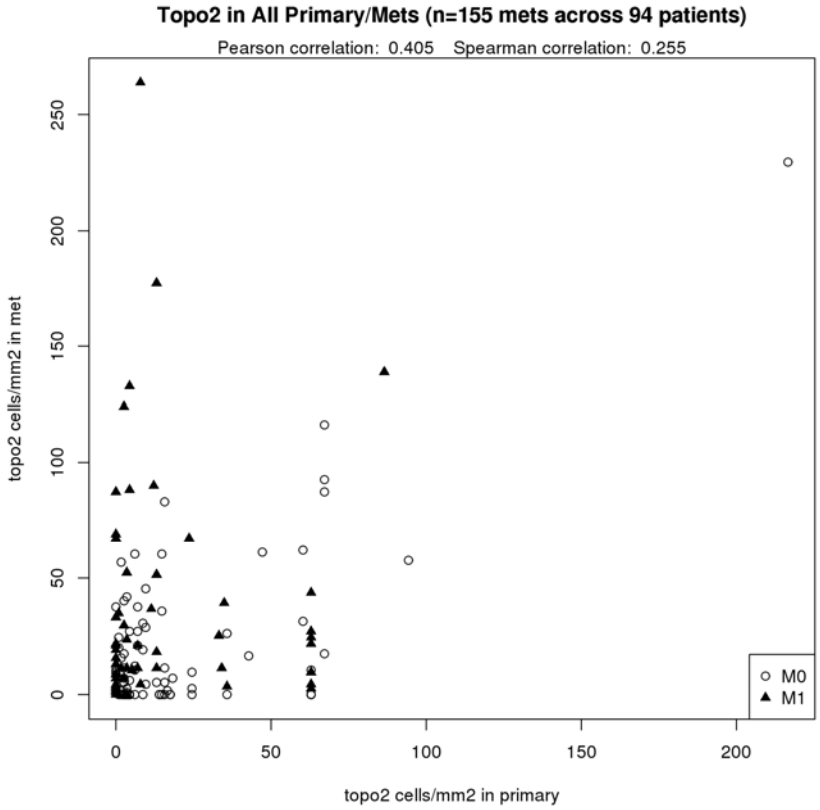
clinic	timing	Site	Metastasectomy.Date	BAP1.Pri	BAP1.Met
6139526	M1	4=BONE	12/1/2003	p	n
6139526	M1	4=BONE	8/25/2004	p	p

Correlation in TOPOII

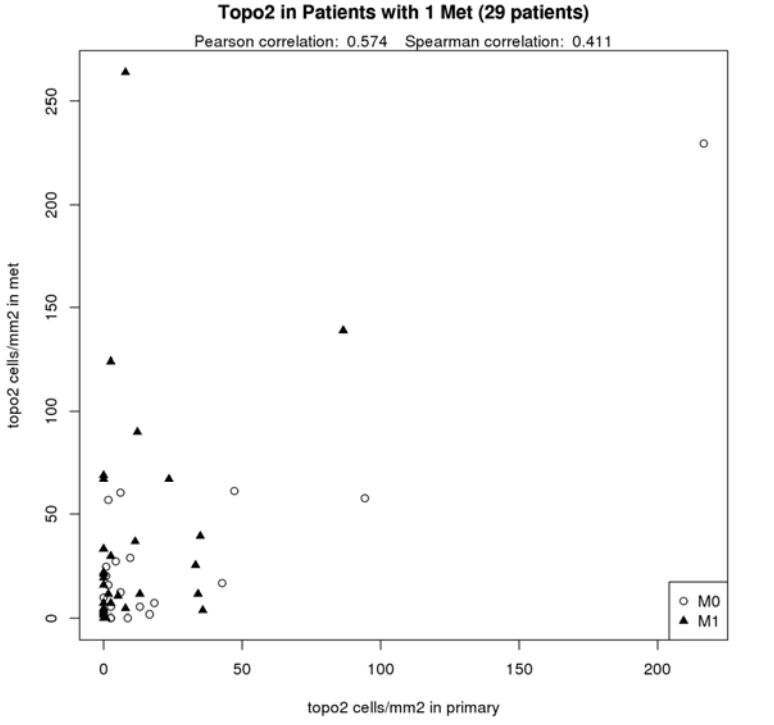


*Minus the outlier above: Pearson=0.426, Spearman=0.156





*Minus the M0 outlier above: Pearson=0.217, Spearman=0.241



*Minus the M0 outlier above: Pearson=0.303, Spearman=0.379

TOPOII Primary/Met Correlation by Met Site

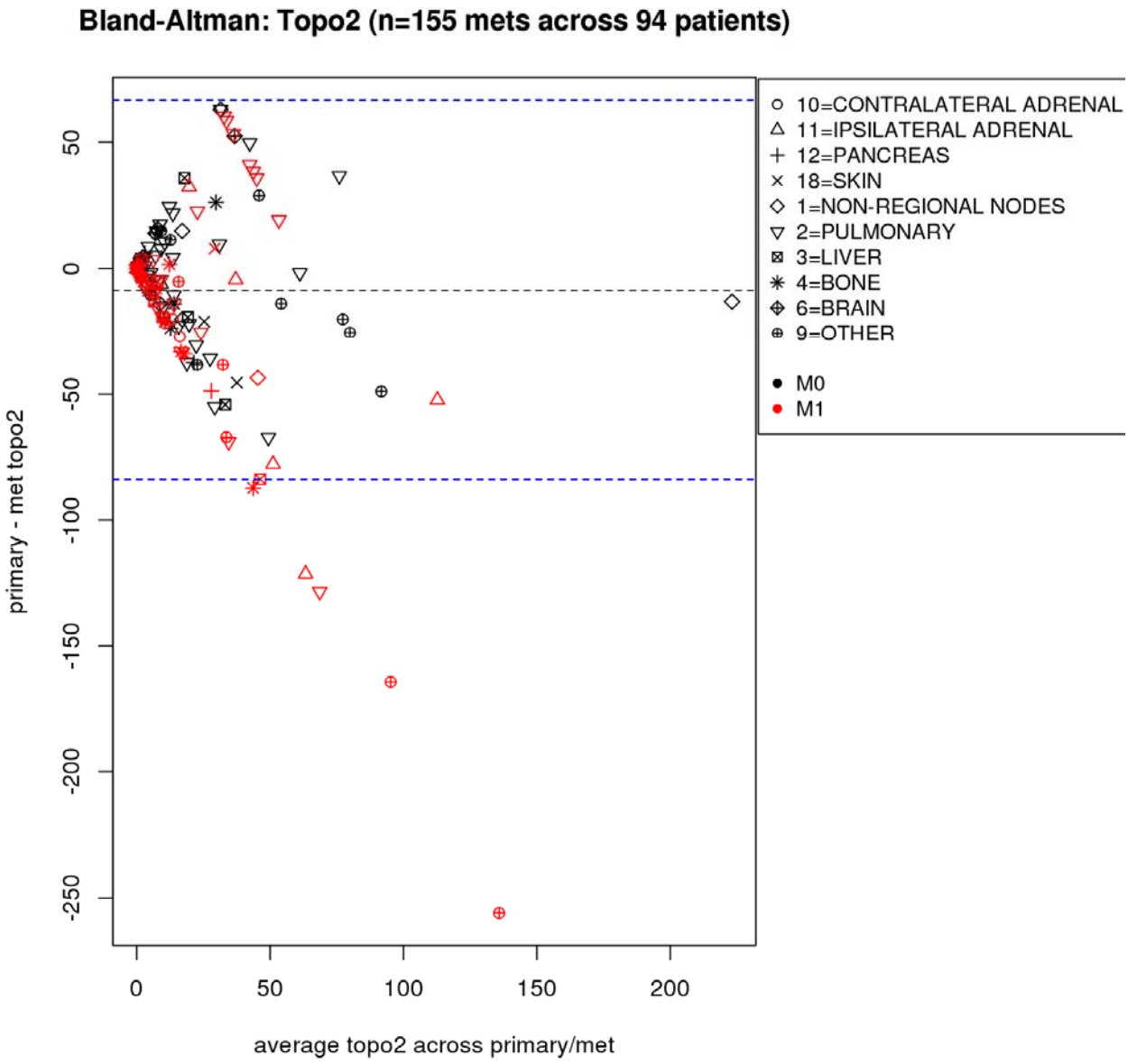
	N	Pearson Corr	Spearman Corr	Median Topo2
CONTRALATERAL ADRENAL	6	-0.23	0	5.2
IPSILATERAL ADRENAL	8	0.49	0.59	25.8
PANCREAS	5	0.12	-0.05	9.6
THYROID	1	---	---	1.7
SKIN	4	0.35	0.32	30.5
BOWEL	1	---	---	10.5
NON-REGIONAL NODES	7	0.92	0.52	9.6
SPLEEN	2	---	---	101.8
MUSCLE	1	---	---	0
OMENTUM	1	---	---	51.5
PULMONARY	69	0.15	0.18	6.1
LIVER	9	-0.19	0.06	13.1
BONE	18	-0.04	-0.08	10.1
HEART	1	---	---	263.8
BRAIN	9	0.31	-0.17	1.7
OTHER	13	0.31	0.5	21.8

	M0				M1			
	N	Pearson Corr	Spearman Corr	Median Topo2	N	Pearson Corr	Spearman Corr	Median Topo2
CONTRALATERAL ADRENAL	2	-1	-1	3.5	4	0.93	0.74	13.6
IPSILATERAL ADRENAL	1	---	---	12.2	7	0.46	0.64	39.3
PANCREAS	3	-0.31	-0.5	9.6	2	1	1	28.4
THYROID	1	---	---	1.7	0	---	---	---
SKIN	2	---	---	48	2	1	1	12.7
BOWEL	1	---	---	10.5	0	---	---	---
NON-REGIONAL NODES	6	0.95	0.58	4.8	1	---	---	67.2
SPLEEN	2	---	---	101.8	0	---	---	---
MUSCLE	0	---	---	---	1	---	---	0
OMENTUM	0	---	---	---	1	---	---	51.5
PULMONARY	42	0.28	0.05	5.2	27	0.01	0.38	11.4
LIVER	3	-0.9	-1	28.8	6	0.37	0.03	11.8
BONE	8	0.32	0.06	5.6	10	-0.17	-0.03	11.4
HEART	0	---	---	---	1	---	---	263.8
BRAIN	5	0.24	-0.22	0	4	-0.57	-0.82	1.7
OTHER	6	0.65	0.6	36.6	7	0.6	0.47	18.3

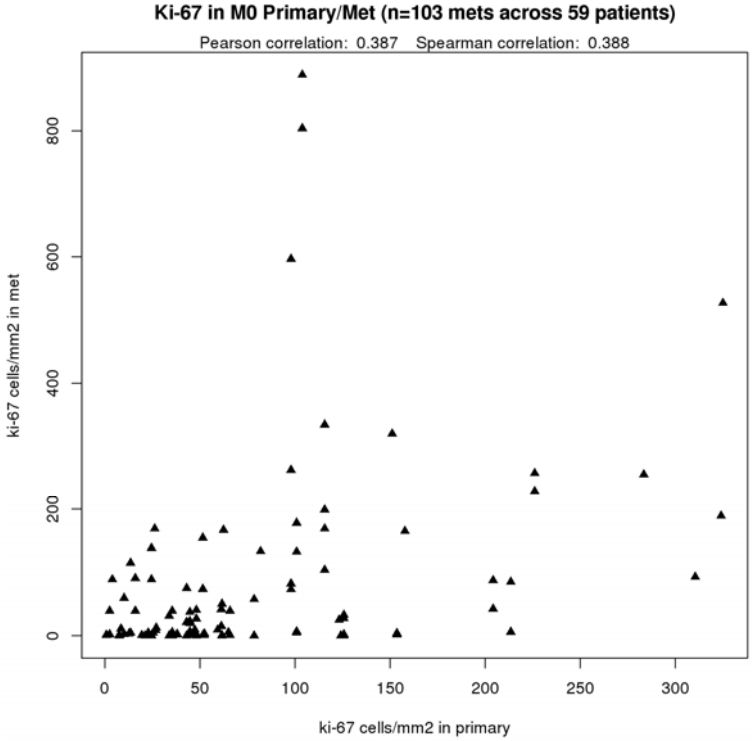
Difference between metastases within patient

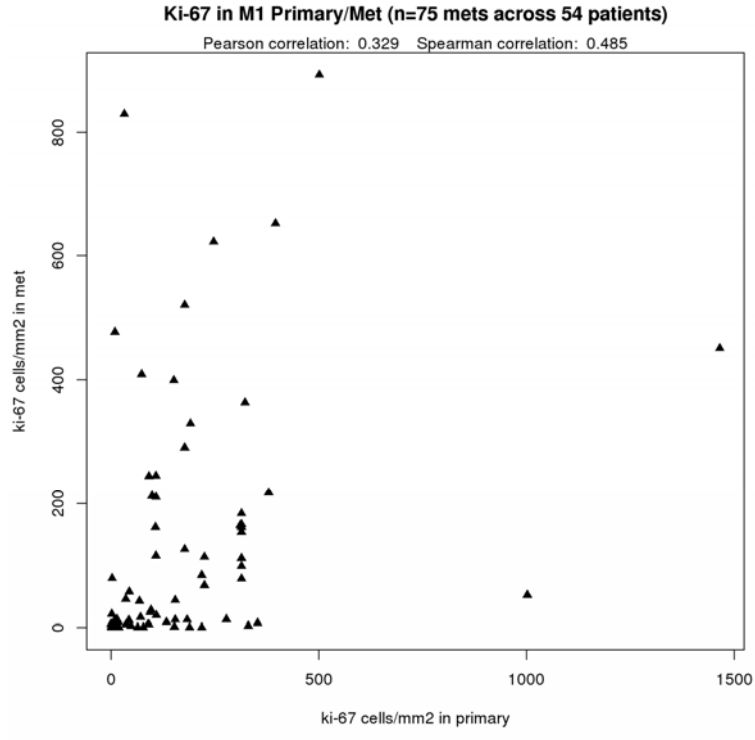
For patients with more than one met stained (n=36 patients), I've looked at the mean absolute deviation (MAD) amongst their mets. As a reminder, MAD is the average distance from the mean and thus an intuitive measure of precision. The table of MADs is below; the average per-patient MAD is 11.6. Based on this I think it's safe to say that Topo2 expression is similar within most patient's mets, a few high-MAD cases not-withstanding.

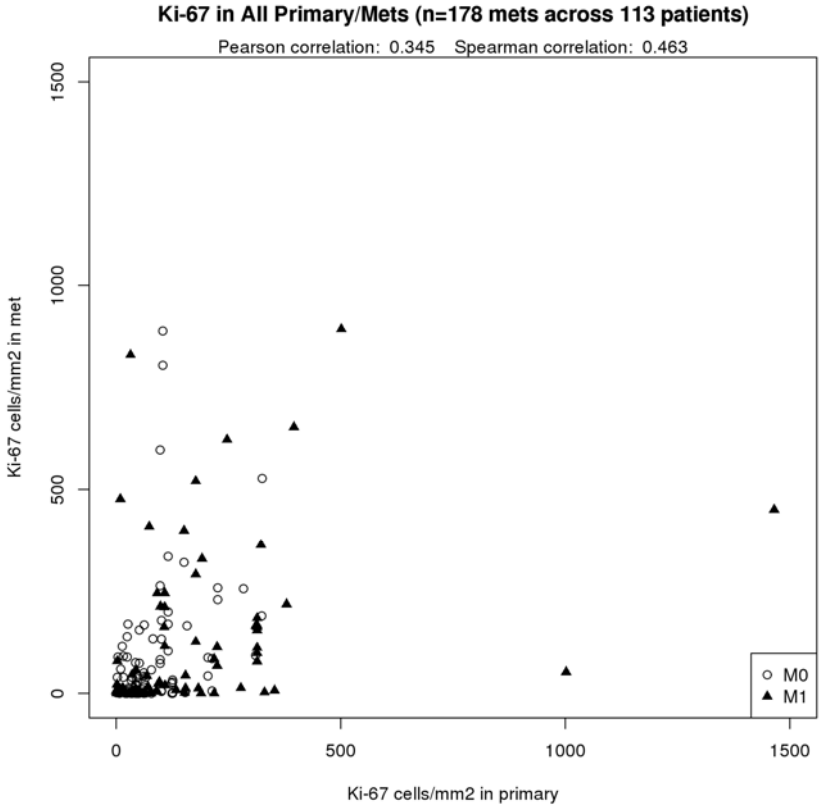
Timing	# Mets	MAD	TOPOII
M1	3	63.3	18.3 177.3 51.5
M1	2	39.3	87.3 8.7
M0	4	30.4	17.5 116.2 92.6 87.3
M0	4	29.1	83 0 5.2 11.4
M1	2	22.3	88.2 132.8
M0	3	21.4	35.8 0 60.3
M0	2	20.5	45.4 4.4
M1	2	17.4	34.9 0
M0	2	16.6	40.2 7
M1	2	16.6	0 33.2
M0	4	15.3	41.9 0.9 2.6 0
M0	2	15.3	62 31.4
M1	2	14.4	52.4 23.6
M0	2	13.6	37.6 10.5
M0	2	13.1	0 26.2
M1	7	11.6	24.5 4.4 21.8 27.1 2.6 43.7 9.6
M0	3	6	27.1 21 37.6
M1	3	6	19.2 13.1 2.6
M0	4	5.9	0 17.5 5.2 11.4
M0	2	5.7	30.6 19.2
M1	2	5.7	11.4 0
M0	2	5.2	0 10.5
M1	2	4.8	21 11.4
M0	4	3.9	10.5 0 0 0
M0	3	3.7	2.6 9.6 0
M0	2	3.5	7 0
M0	2	3	8.7 2.6
M1	2	1.8	0 3.5
M0	7	1.5	0 6.1 0 0 0 0 0
M0	2	0.8	1.7 0
M1	2	0.4	1.7 0.9
M0	2	0	0 0
M0	2	0	0 0
M0	2	0	0 0
M1	2	0	0 0
M1	2	0	0 0

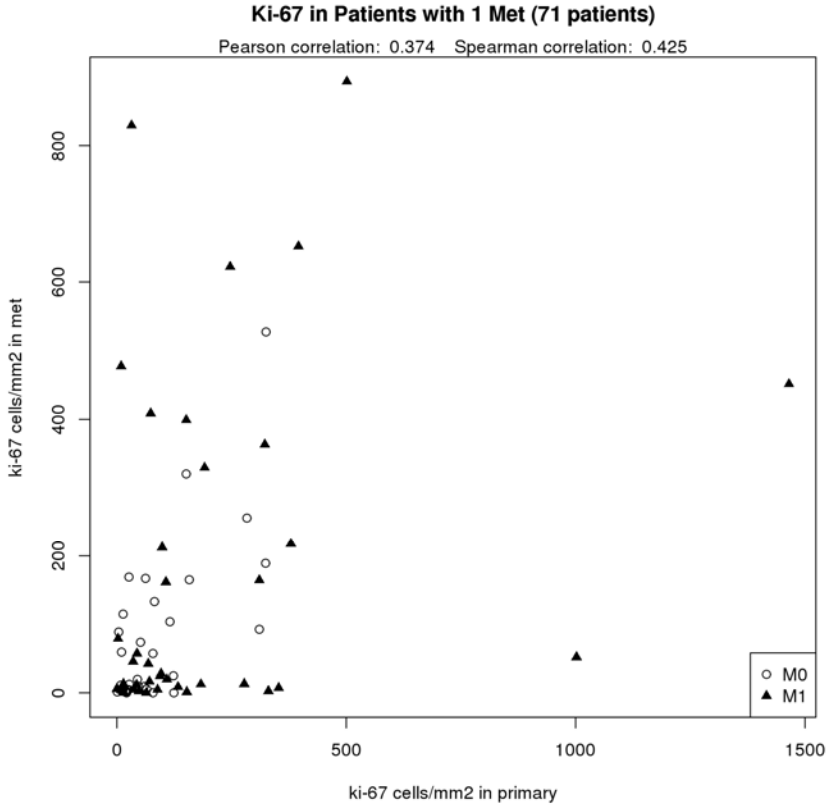

Bland-Altman: Topo2 (n=155 mets across 94 patients)
primary - met topo2
50
0
-50
-100
-150
-200
-250
0
50
100
150
200
average topo2 across primary/met
10=CONTRALATERAL ADRENAL
11=IPSILATERAL ADRENAL
12=PANCREAS
18=SKIN
1=NON-REGIONAL NODES
2=PULMONARY
3=LIVER
4=BONE
6=BRAIN
9=OTHER
M0
M1

Correlation in Ki-67






Ki-67 in All Primary/Mets (n=178 mets across 113 patients)
Pearson correlation: 0.345 Spearman correlation: 0.463
Ki-67 cells/mm2 in met
Ki-67 cells/mm2 in primary
0
500
1000
1500
M0
M1


Ki-67 in Patients with 1 Met (71 patients)
Pearson correlation: 0.374 Spearman correlation: 0.425
ki-67 cells/mm2 in met
ki-67 cells/mm2 in primary
0
200
400
600
800
500
1000
1500
M0
M1

Ki-67 Primary/Met Correlation by Met Site

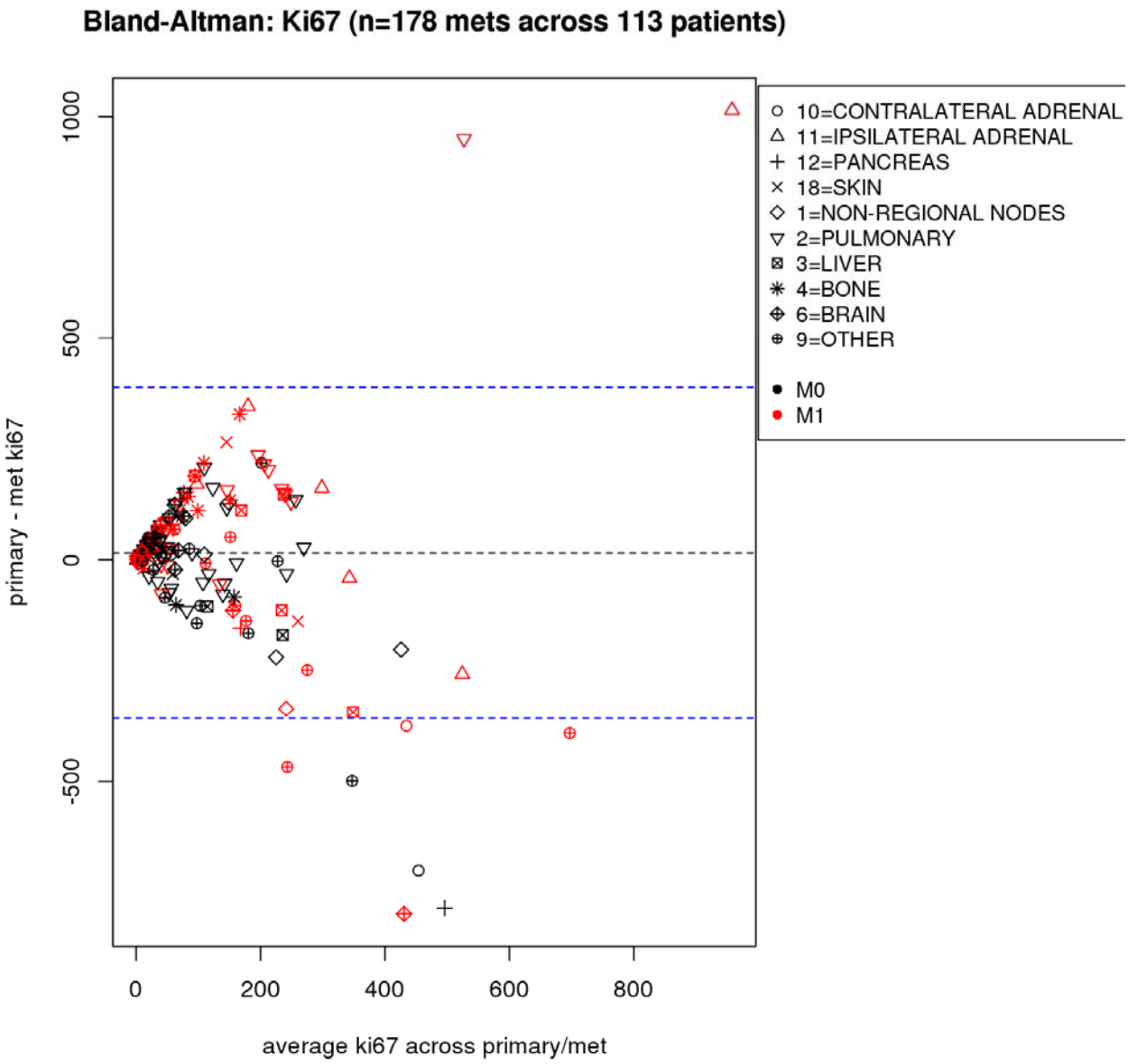
	N	Pearson Corr	Spearman Corr	Median Ki-67
CONTRALATERAL ADRENAL	8	0.66	0.83	12.1
IPSILATERAL ADRENAL	10	0.61	0.89	9.8
PANCREAS	8	0.38	0.43	2.3
THYROID	1	---	---	4.6
SKIN	4	0.19	-0.32	48.3
BOWEL	1	---	---	39.1
NON-REGIONAL NODES	8	0.54	0.47	94.2
SPLEEN	2	---	---	334.9
MUSCLE	1	---	---	0
OMENTUM	1	---	---	210.6
PULMONARY	77	0.45	0.47	13
LIVER	11	0.31	0.11	113.5
BONE	19	0.06	0.04	3.9
HEART	1	---	---	399.1
BRAIN	11	-0.22	-0.02	5.9
OTHER	15	0.65	0.46	125.9

	M0				M1			
	N	Pearson Corr	Spearman Corr	Median Ki-67	N	Pearson Corr	Spearman Corr	Median Ki-67
CONTRALATERAL ADRENAL	3	0.95	1	50.2	5	0.91	0.97	4.6
IPSILATERAL ADRENAL	2	1	1	3.3	8	0.55	0.85	115.1
PANCREAS	6	0.77	0.43	2	2	-1	-1	122.9
THYROID	1	---	---	4.6	0	---	---	---
SKIN	2	---	---	48.3	2	-1	-1	171.2
BOWEL	1	---	---	39.1	0	---	---	---
NON-REGIONAL NODES	7	0.74	0.59	84.8	1	---	---	408.9
SPLEEN	2	---	---	334.9	0	---	---	---
MUSCLE	0	---	---	---	1	---	---	0
OMENTUM	0	---	---	---	1	---	---	210.6
PULMONARY	51	0.61	0.3	8.5	26	0.46	0.71	44
LIVER	5	0.34	-0.1	39.1	6	0.18	-0.06	138.9
BONE	9	0.4	0.32	3.3	10	0.14	-0.12	8.2
HEART	0	---	---	---	1	---	---	399.1
BRAIN	7	-0.06	0.16	4.6	4	-0.14	0	109.2
OTHER	7	0.24	0.54	154.5	8	0.76	0.58	120.6

Difference between metastases within patient

For patients with more than one met stained (n=40 patients), I've looked at the mean absolute deviation (MAD) amongst their mets. The table of MADs is below; the average per-patient MAD is 29.3. Clearly variation amongst Ki-67 in mets is higher than for TOPOII.

Timing	# Mets	MAD	Ki-67
M0	4	176.1	82.16 262.8 73.04 596.67
M1	3	138.7	520.38 290.84 125.86
M1	2	120	4.56 244.54
M0	4	74.8	178.02 6.52 132.38 4.56
M0	3	68.4	0.65 0.65 154.55
M0	3	66.9	334.53 168.89 198.89
M1	3	50	245.19 115.42 210.63
M0	2	42.1	804.04 888.16
M1	2	42.1	0 84.12
M0	2	39.5	84.77 5.87
			111.51 183.89 161.72 78.25 98.47 153.9
M1	7	34.5	166.29
M0	3	28.5	0 74.99 21.52
M0	2	25.8	90.64 39.13
M0	2	24.8	138.25 88.69
M1	2	22.8	113.47 67.82
M0	2	22.5	87.38 42.39
M0	3	22.2	50.21 0.65 0
M0	2	19.2	0.65 39.13
M0	2	18.6	1.96 39.13
M0	2	17	5.22 39.13
M1	2	15.7	43.69 12.39
M0	4	14.5	1.96 32.61 0 27.39
M0	2	14.3	257.58 228.89
M0	3	13.9	0 31.3 0
M0	7	13.5	40.43 0 1.96 0 0 0 26.74
M0	2	13.4	41.73 15
M0	4	12.9	1.96 37.17 6.52 22.82
M1	2	10.1	21.52 1.3
M1	2	4.2	11.09 2.61
M0	2	1.3	0.65 3.26
M0	2	1	1.96 3.91
M0	2	0.7	1.3 0
M0	2	0.7	1.3 0
M0	2	0.3	0 0.65
M0	2	0.3	2.61 1.96
M1	2	0	0 0
M1	2	0	0 0
M1	2	0	0 0
M1	2	0	0 0
M1	2	0	1.3 1.3



Summary

Our work in this task has 1) validated our BioPROG score and 2) advanced our knowledge of the dual expression of key biomarkers in paired sets of primary and metastatic ccRCC. We are leveraging the latter now for publications and expansion in to new drug development.

Task 7: See Publications, Abstracts and Presentations section

Key Research Accomplishments:

1. Successful MALDI-MS imaging analysis of 68 renal tissue samples obtained from the Mayo Clinic, and an additional 20 tissues with tumor/margin/non-tumor regions from Eastern Virginia Medical School (via Dean Troyer, M.D.).

2. Established an optimized frozen ccRCC tissue preparation protocol with Mayo Clinic suitable for MALDI MS imaging and metabolomic analyses.

3. Identification of 38 candidate m./z protein peaks that are differentially expressed in recurrent ccRCC or non-recurrent ccRCC tissues.

4. Identification of thymosin beta family peptides and S100 family member proteins as being differentially expressed in recurrent ccRCC tumors.

5. Established MALDI MS imaging methods to effectively profile the distribution of N-linked glycans, ceramides, sphingomyelins, phospholipids and small molecule metabolites.

6. From the novel MALDI MS imaging experiments, a unique database of N-glycan, sphingolipid and glycosphingolipid species has been created in the Drake laboratory. This is being distributed via publications, or freely available upon request.

7. Developed a clinical scoring algorithm (BioPROG) that is being developed as a CLIA-approved test to help urologic surgeons provide better, more individualized management following surgery for localized ccRCC.

Conclusion:

The studies reinforce the need to look at all molecular analytes in the analysis of biomarkers of ccRCC. The analysis of lipids, metabolites and glycans directly in ccRCC tissues balance the known heterogeneity of ccRCC at the genomic level, and allow a more complete assessment of the properties of ccRCC than can be provided by genetic, epigenetic and/or proteomic studies alone. The methods developed from this grant support are being applied to increased numbers of ccRCC tissues, with emphasis on tissues with bisecting margin areas. Glycan analysis of stage specific tissue microarrays is also ongoing. Being able to spatially assess the tissue distribution of potential biomarker analytes is a critical new approach to be applied with existing biomarker discovery tools. Finally, our work also demonstrates that while some biomarkers appear to be carried forward from primary to metastatic ccRCC, we also note that in some cases there is considerable variability (loss and gain of expression) between primary and metastatic ccRCC. This is informative for the field as it further underscores the need to take in to account the clonal expansion theory of ccRCC progression and also to move away form a drug development approaches that focus solely on observations of potential targets made in primary ccRCCs (as these targets may not be carried forward to metastatic ccRCC and more importantly new targets may develop in metastatic ccRCC as well).

Publications, Abstracts and Presentations:

Publications*:

1. Jones, E.E., Powers, T.W., Neely, B.A., Cazares, L.H., Troyer, D.A., Parker, A.S., and Drake, R.R. (2014) MALDI Imaging Mass Spectrometry Profiling of Proteins and Lipids in Clear Cell Renal Cell Carcinoma. Proteomics, 14, 924-935. PMID: 24497498

2. Powers, T.W., Jones, E.E., Betesh, L.R., Romano, P., Gao, P., Copland, J.A., Mehta, A.S., Drake, R.R. (2013) A MALDI Imaging Mass Spectrometry Workflow for Spatial Profiling Analysis of N-linked Glycan Expression in Tissues. Anal. Chem., 85, 9799-9806. PMID: 24050758.

*PDF copies of these publications are attached in the Appendix

Related Publications in kidney tissues resulting from methods developed for this grant:

3. Powers, T.W., Neely, B.A., Shao, Y., Tang, H., Troyer, D.A., Mehta, A.S., Haab, B.B., and Drake, R.R. (2014) MALDI Imaging Mass Spectrometry Profiling of N-Glycans in Formalin-Fixed Paraffin Embedded Clinical Tissue Blocks and Tissue Microarrays. PLoS One, In press

4. Nowling, T.K., Mather, A.R., Thiyagarajan, T., Hernandex-Corbacho, M.J., Jones, E, Powers, T.W., Snider, A., Oates, J.C., Drake, R.R., and Siskind, L.J. (2014) Renal glycosphingolipid metabolism is dysfunctional in lupus mice and patients with nephritis. J. Am. Soc. Nephrol., In press.

5. Jones, E.E., Dworski, S., Canals, D., Casas, J., Fabrias, G., Schoenling, D., Levade, T., Denlinger, C., Hannun, Y.A., Medin, J.A., and Drake, R.R. (2014) On-Tissue Localization of Ceramides and other Sphingolipids by MALDI Mass Spectrometry Imaging. Anal. Chem. 86: 8303-8311.

6. Parker AS**,** Eckel-Passow JE, Serie D, Hilton T, Parasramka M, Joseph RW, Wu KJ, Cheville JC, Leibovich BC. Higher Expression of Topoisomerase II Alpha Is an Independent Marker of Increased Risk of Cancer-specific Death in Patients with Clear Cell Renal Cell Carcinoma. Eur Urol. 2013 Dec 25. PMID:24388441. DOI:10.1016/j.eururo.2013.12.017.

7. Eckel-Passow JE, Serie DJ, Bot BM, Joseph RW, Hart SN, Cheville JC, Parker AS. Somatic expression of ENRAGE is associated with obesity status among patients with clear cell renal cell carcinoma. Carcinogenesis. 2014 Apr; 35(4):822-7. Epub 2013 Dec 28. PMID:24374825. PMCID:3977147. DOI:10.1093/carcin/bgt485.

8. Joseph RW, Kapur P, Serie DJ, Eckel-Passow JE, Parasramka M, Ho T, Cheville JC, Frenkel E, Rakheja D, Brugarolas J, **Parker A**. Loss of BAP1 protein expression is an independent marker of poor prognosis in patients with low-risk clear cell renal cell carcinoma. Cancer. 2014 Apr 1; 120(7):1059-67. Epub 2013 Dec 30. PMID:24382589. DOI:10.1002/cncr.28521.

9. Eckel-Passow JE, Igel DA, Serie DJ, Joseph RW, Ho TH, Cheville JC, Parker AS. Assessing the clinical use of clear cell renal cell carcinoma molecular subtypes identified by RNA expression analysis. Urol Oncol. 2014 Aug 27. pii: S1078-1439(14)00276-2. doi: 10.1016/j.urolonc.2014.07.019. [Epub ahead of print] PMID: 25175426.

10. Ho TH, Kapur P, Joseph RW, Serie DJ, Eckel-Passow JE, Parasramka M, Cheville JC, Wu KJ, Frenkel E, Rakheja D, Stefanius K, Brugarolas J, Parker AS. Loss of PBRM1 and BAP1 expression is less common in non-clear cell renal cell carcinoma than in clear cell renal cell carcinoma. Urol Oncol. 2015 Jan;33(1):23.e9-23.e14. doi: 10.1016/j.urolonc.2014.10.014. Epub 2014 Nov 24. PMID: 25465300

Oral presentations:

"Protein and Lipid Profiling of Recurrent and Non-Recurrent Renal Cell Carcinoma" Ourense Conference on Imaging Mass Spectrometry, Ourense, Spain, Sept. 2012

"MALDI Mass Spectrometry Imaging for On-Tissue Spatial Profiling of N-linked Glycans in Tumor Tissues." Mass Spectrometry Applications to the Clinical Laboratory Annual Meeting, San Diego, CA, February 2013.

"Differential Molecular Profiling of Lipids and Glycans at the Tumor Margin of Clear Cell Renal Carcinoma Tissues by MALDI-MS Imaging." American Society of Mass Spectrometry Annual Meeting, Minneapolis, MN, June 2013.

"Novel on-tissue glycosidase and lipase digestion workflows to identify complex glycan and lipid species by MALDI-TOF MS imaging." Bruker Daltonics User Meeting, one of 9 invited speakers, ASMS Annual Meeting, Minneapolis, MN, June 2013.

Poster presentations:

"Defining the Molecular Tumor Margin Regions of Clear Cell Renal Carcinoma Tissues by MALDI-MS Imaging of Lipid and Glycan Species." Genetourinary Cancer Symposia, American Society of Clinical Oncology, Orlando, FL, February 2013.

"MALDI mass spectrometry imaging for on-tissue spatial profiling of tumor-specific N-linked glycans and lipids in renal carcinomas." American Association of Cancer Research Annual Meeting, Washington, DC, April 2013.

"Molecular Characterization of the Clear Cell Renal Carcinoma Margin Interface by MALDI Tissue Imaging of Glycans, Lipids and Proteins." Kidney Cancer Foundation, Dublin, Ireland, March 2014.

Conclusions: We have developed a combined proteomics, metabolomics, lipidomic and glycomic workflow to analyze recurrent and non-recurrent RCC tissues. We expect to submit multiple publications derived from this research, as well as several new grant applications.

Inventions, Patents and Licenses:

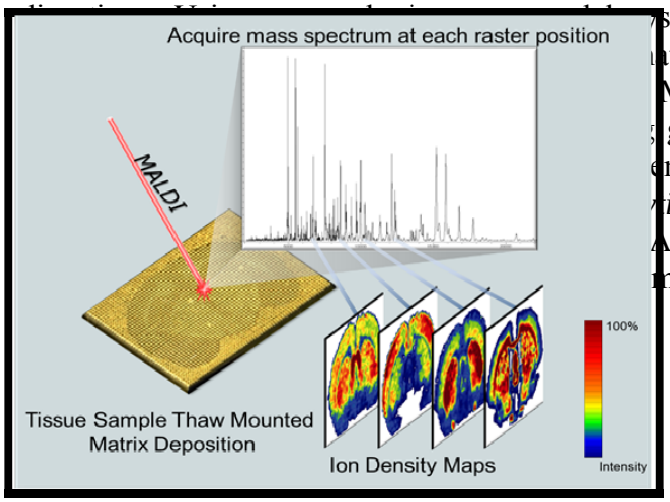
None

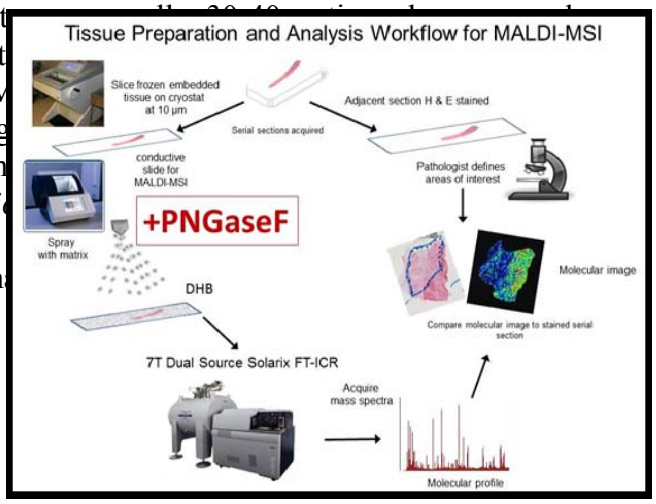
Reportable Outcomes:

An entirely new MALDI MS imaging method has been developed for on-tissue analysis of the distribution of N-glycans. This approach has major implications for localization and function of specific glycans in tumor tissues and surrounding stroma and other sub-structures.

Appendix Material:

Our group has been recently optimized and described a tissue profiling MALDI-IMS approach to uniquely profile the expression and distribution of N-glycans released by on-tissue PNGaseF





MALDI Imaging of N-Glycans and Comparison to Histopathology. ccRCC tissues were cut at 10 µm and sequentially ethanol washed to remove lipids. PNGase F (20 mU) was applied using the Bruker ImagePrep, followed by a 2 Hr incubation at 37ºC. DHB matrix was then applied, prior to imaging. As a control, duplicate ccRCC tissues were analyzed with and without PNGase F treatment, then extracted in water for off-tissue profiling (**Fig 1A.**). A comparison of the mass spectra reveals robust differences between the PNGase F treated and non treated tissues, and further this is further illustrated in the glycan profile images linked to the H&E overlay for three representative glycans at m/z 2486, 2631 and 2059 (**Fig 1B.**). Each individual mass to charge ratio (m/z) can be queried for peak intensities in each spot analyzed. A color pixel scale is used to convert the intensities to a colar representation; hence an image of color intensity is generated for each m/z of interest.

Figure 1A.

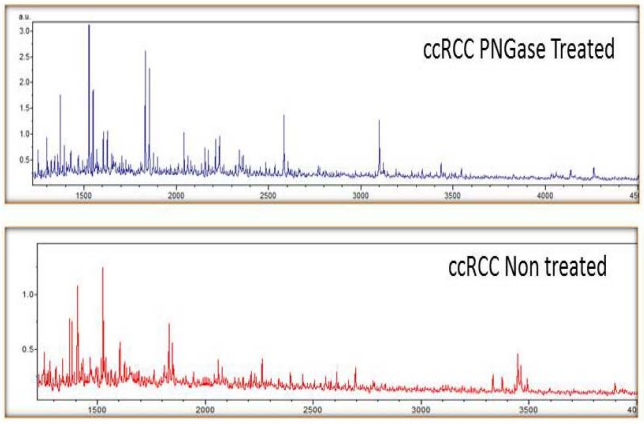
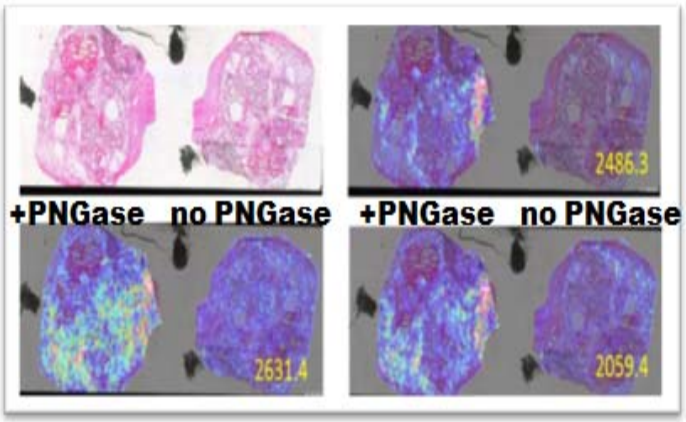


Figure 1B.



In the course of these initial analyses, it was observed that there were specific examples of glycans that were expressed in the margin region between tumor and non-tumor regions. This was further analyzed using larger tissue specimens specifically collected for that purpose, provided by Dr Dean Troyer, Eastern Virginia Medical School, and eight samples have already been analyzed for on-tissue PNGaseF digestions and MALDI-IMS. Representative profiles of one tissue is shown in the two panels below (Figure 2A/2B).

Figure 2A.

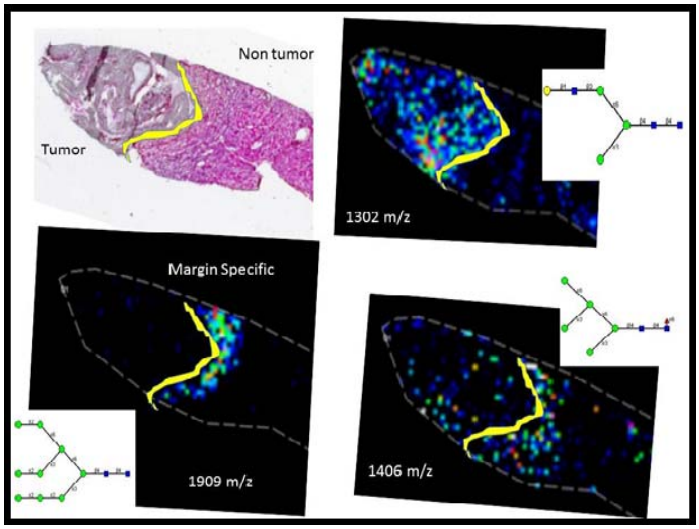
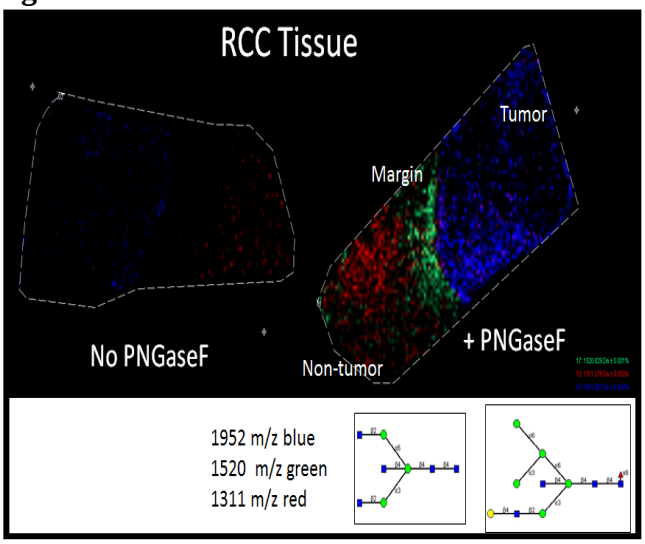


Figure 2B.



There are distinct glycosylation profiles between tumor, non-tumor and margin regions adjacent to the tumor. The yellow band along the margin was added to illustrate the boundary of the region.We have made glycan structure determinations based on accurate mass, databases and off-tissue analysis. Two examples of structural assignments following PNGase F on-tissue digest of an RCC tissue is shown in Figure 3A/3B. For both structures below, searches of the native mass CFG database, and comparison to permethylated glycans of extracted glycans from the same tissue. These two species represent distinct glycan species associated with expression, or lack thereof, at the margin interface of tumor and non-tumor regions.

Figure 3A

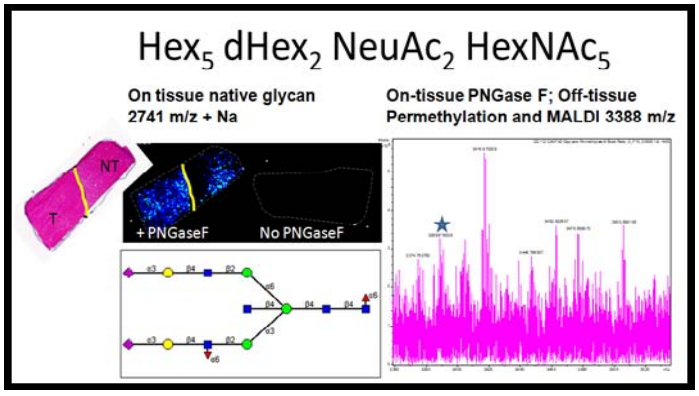
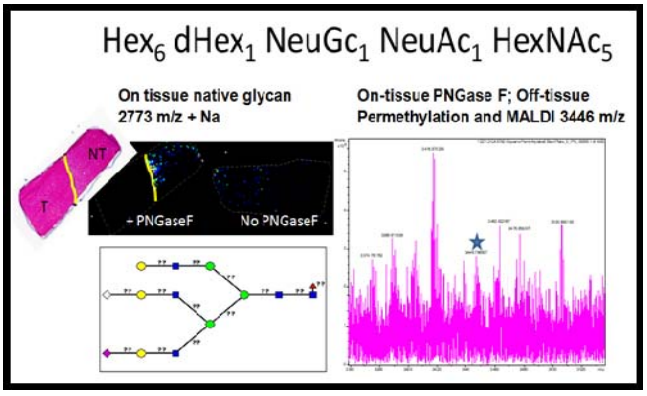


Figure 3B



Preliminary Protein Analysis of Margin vs. Tumor and non-tumor regions. Because the margin glycan expression was so well defined, 1 mm slices were scraped from the margin region, and a similar scrape from tumor and non-tumor regions. Tissues were digested in TFE and digested with tryspin, and analyzed by LC-MS for total protein composition. The Venn diagram (Figure 4A) summarizes the total number of proteins identified from each tissue region,

and the table highlights several extracellular matrix proteins identified exclusively in the margin region (out of 228)(Figure 4B). This region is clearly distinct from the other two regions analyzed, and reflects protein differences consistent with an active EMT process. This baseline data will be used with the glycopeptide analyses described in Figure 5.

Figure 4A

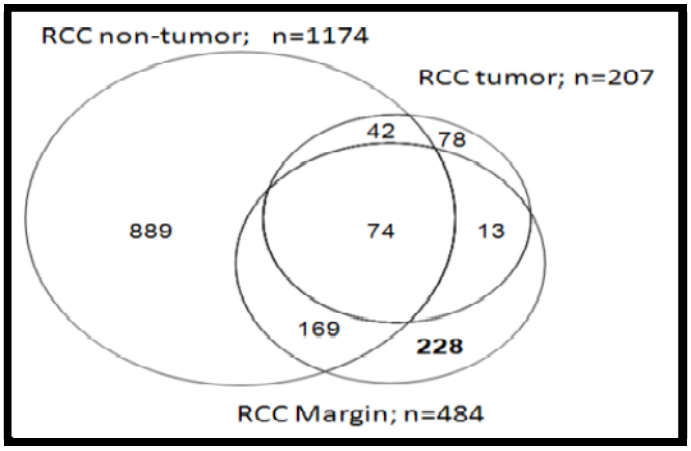
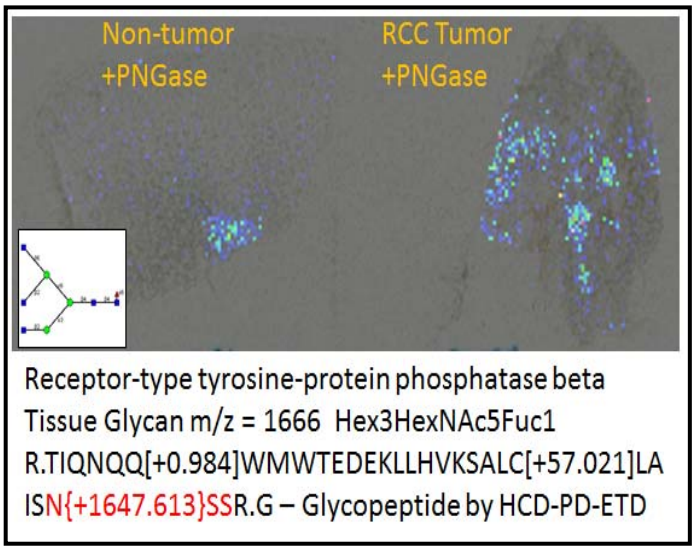


Figure 4B

Protein Name	# Unique Peptides	MASCOT score
Extracellular Matrix/ Adhesion/Junctions		
Protocadherin 17	3	56.5
Protocadherin 20	2	58.4
ADAMTS18	4	59.2
ADAMTS4	2	50.7
Integrin alpha-3	3	55.3
Neural Cell Adhesion Molecule L1	3	55.4
Laminin subunit gamma 3*	3	70.5
Fer like protein 5*	2	51.9
Synaptajanin-2	2	87.3
Vitronectin	2	50.2
Cubulin	3	79.4
Neurexin-1 beta*	3	54.9

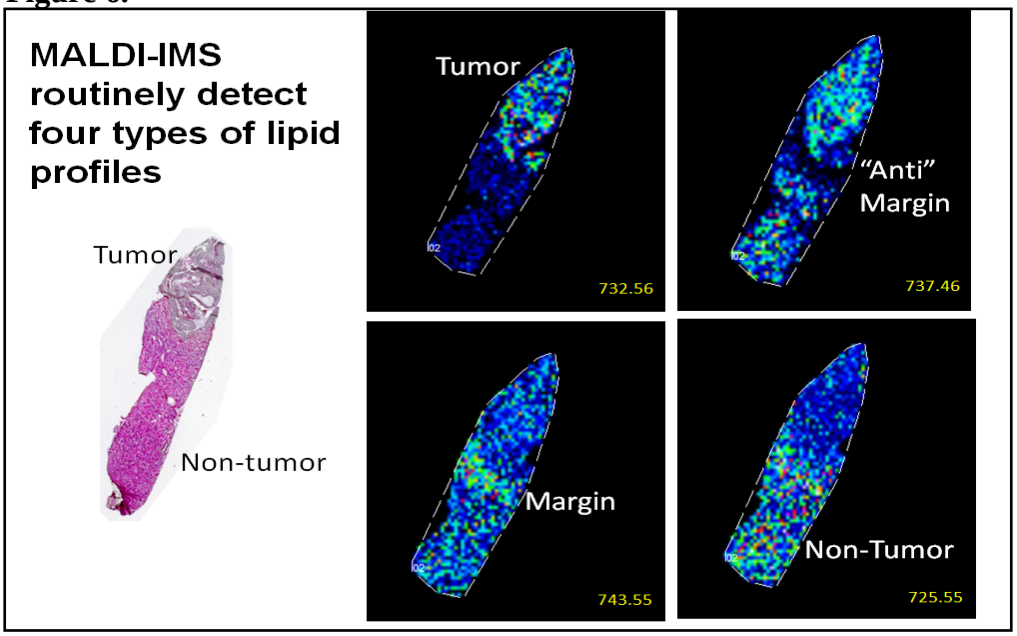
HCD-PD-ETD Glycopeptide from RCC Tissue Example: In the past, MS-based glycoproteomic studies relied on the release of glycan moieties from glycopeptides followed by separate MS analysis of the glycans and the peptides. Recently, intact glycopeptide strategies are emerging for use with high resolution Orbitrap instruments, typified by Higher-energy Collision Dissociation-Product Dependent-Electron Transfer Dissociation (HCD-PD-ETD). This method is a data-dependent acquisition, based on detection of glycan oxonium ions that trigger further analysis of the peptide carrier by ETD. An initial analysis of glycopeptides from the tumor region of one RCC tissue has been done, using a HILIC resin enrichment and HCD-PD-ETD protocol Glycopeptide data was obtained on a Thermo Orbitrap Elite mass spectrometer. In Figure 5 (below), a glycopeptide of receptor-type protein phosphastase beta was identified. The same glycan ion could also be detected in the glycan MALDI imaging profile of the same tumor tissue.

Figure 5.



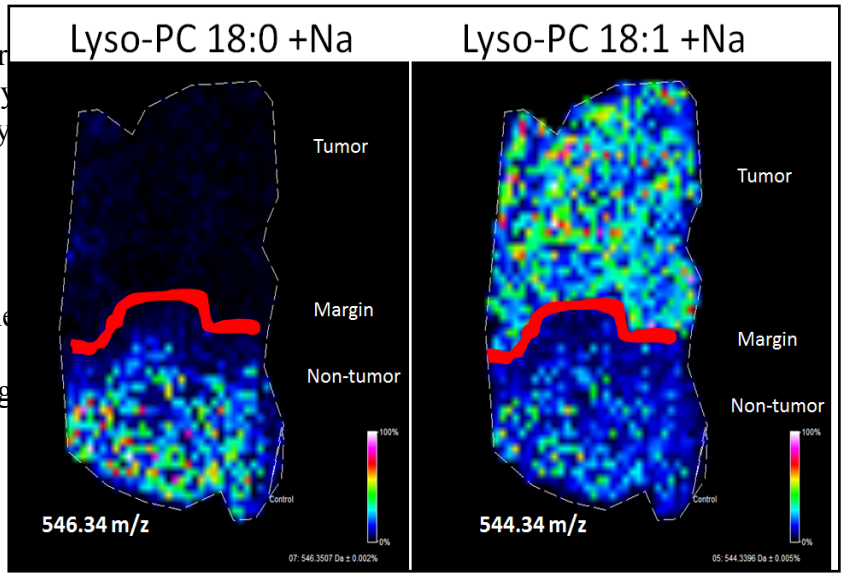
MALDI Imaging of Lipids and Comparison to Histopathology. ccRCC tissues (n=17) were cut at 10 µm, washed in water, dessicated and sprayed with DHB using an ImagePrep. Lipid images were obtained on a 7T Dual Source Solarix FT-ICR Mass spectrometer (Bruker Daltonics). Phosphatidylcholines, lyso-PCs, sphingomyelins and ceramides were profiled in positive ion mode. The image panel below in Figure 6 illustrates the type of distribution profiles that are obtained.

Figure 6.



Lyso-PCs and other lipids with 1 or 2 double bonds in the fatty acid chains are associated with the presence of RCC. An example image of a Lyso-PC (C18:0) in non tumor tissues compared to a Lyso-PC (C18:1) is shown below in Figure 7.

Figure 7.



In the data summary shown in Figur 8 below, 30 lipid species detected by MALDI-imaging mass spectrometry in 17 RCC tissues are listed. Lipids differentially expressed in either tumor or normal in at least 14 of 17 sample pairs are listed. Representative images for two of the lipid species for normal or tumor expression are shown in the imaging panels on the right, with an overlay of the two expression profiles also included.

Figure 8.

m/z	Elevated in Normal
627.54	17
625.52	16
916.52	16
575.51	15
599.51	15
720.59	15
724.53	15
731.61	15
769.57	15
780.56	15
790.54	15
806.57	15
672.42	14
753.59	14
778.48	14
828.56	14

m/z	Elevated in Tumor
671.58	16
763.47	16
718.58	15
810.61	15
812.62	15
848.56	15
850.58	15
492.1	14
582.3	14
663.55	14
786.61	14
794.61	14
834.6	14
874.58	14

Out of 17 ccRCC tissues analyzed